



OPEN

## A thinking innovation strategy based Northern goshawk optimizer enhanced extreme learning machine for bankruptcy prediction problems

Keyu Jiang<sup>1,7</sup>, Xuhai Zhao<sup>2,3,7</sup>✉, Yulin Li<sup>4,5</sup> & Baiqiang Chen<sup>6</sup>

Bankruptcy risk prediction, a core issue in financial risk management, plays a critical role in assessing corporate financial health and supporting decision-making for financial institutions. Traditional machine learning models often struggle with parameter optimization when handling high-dimensional, nonlinear financial data. In contrast, metaheuristic algorithms, owing to their global search capabilities, have emerged as effective tools to enhance model performance. This paper proposes a novel bankruptcy prediction model that integrates a Kernel Extreme Learning Machine (KELM) with an improved Northern goshawk Optimizer (TIS\_NGO), which incorporates a Thought-Inspired Strategy (TIS). The enhancements to TIS\_NGO include a divergence-based thought innovation mechanism, a prey-attacking strategy inspired by differential evolution, and a centroid opposition-based boundary control mechanism. Experimental evaluations on the CEC2017 and CEC2022 benchmark suites demonstrate that TIS\_NGO outperforms the standard NGO as well as other well-known algorithms such as Particle Swarm Optimization (PSO) and Grey Wolf Optimizer (GWO) in terms of convergence speed and solution accuracy. The optimized KELM, with TIS\_NGO-tuned penalty parameter  $C$  and kernel parameter  $\gamma$ , achieves high classification accuracy and robustness on the Wieslaw bankruptcy dataset. These results validate the effectiveness of combining improved metaheuristic algorithms with machine learning models for financial risk forecasting. Overall, the proposed method offers a promising approach to improving the accuracy and stability of bankruptcy prediction, thereby contributing a new technical pathway for early warning in the financial domain.

**Keywords** Kernel extreme learning machine, Northern goshawk optimizer, Parameter optimization, Thinking innovation strategy, Bankruptcy prediction

The escalating global concern over financial crises stems from their profound and far-reaching impacts on enterprises. Against this backdrop, the development of efficient bankruptcy prediction tools has become increasingly critical, as they provide essential decision-making support for businesses<sup>1,2</sup>. Currently, bankruptcy prediction is primarily treated as a binary classification problem, and numerous novel classification models have emerged through the application of diverse technological approaches<sup>3–5</sup>. Existing prediction models can generally be categorized into two types: those based on traditional statistical methods, which feature relatively simple structures, and those grounded in advanced artificial intelligence (AI) techniques. Due to their superior capability in handling nonlinear data, AI-based models often outperform conventional analytical approaches in various performance evaluations<sup>6</sup>. Moreover, the rapid advancement of big data and AI technologies has significantly propelled progress in the field of bankruptcy prediction. In the context of financial decision-making and bankruptcy forecasting, a wide array of AI models has garnered significant attention from both academia

<sup>1</sup>Dundee International Institute of Central South University, Changsha, China. <sup>2</sup>School of Economics and Management, Jiujiang Polytechnic University of Science and Technology, Jiujiang, China. <sup>3</sup>School of Accounting, Guangdong Baiyun University, Guangzhou, China. <sup>4</sup>University of Birmingham, Birmingham, United Kingdom. <sup>5</sup>Jinan University, Guangzhou, China. <sup>6</sup>Graduate School, Sichuan Academy of Social Sciences, Chengdu, China. <sup>7</sup>Keyu Jiang and Xuhai Zhao have contributed equally to this work and should be considered co-first authors. ✉email: zhaoxh@baiyunu.edu.cn

and industry. Representative examples include Artificial Neural Networks (ANN)<sup>7</sup>, Backpropagation Neural Networks (BPNN)<sup>8</sup>, Support Vector Machines (SVM)<sup>9,10</sup>, Decision Trees (DT)<sup>11</sup>, Extreme Learning Machines (ELM)<sup>4</sup>, Kernel Extreme Learning Machines (KELM)<sup>12</sup>, as well as numerous hybrid approaches derived from these models.

Additionally, Bolin Liao et al. employed the Nonlinear Activated Beetle Antennae Search (NABAS) algorithm to address fraud detection in listed companies. The NABAS algorithm demonstrated higher accuracy and efficiency in fraud detection on public datasets compared to other methods<sup>13</sup>. Jabeur et al. proposed an improved XGBoost algorithm based on feature importance selection (FS-XGBoost) for bankruptcy prediction, providing an effective approach for this purpose<sup>14</sup>. Mário Papik conducted one-year, two-year, and three-year predictions across three periods using CatBoost, LightGBM, and XGBoost methods. The results of the study indicate that the predictive models performed significantly weaker during crisis periods compared to non-crisis periods, analyzing the impact of crises on the performance of bankruptcy prediction models<sup>15</sup>. Suguru Yamanaka introduced a novel bankruptcy prediction method that utilizes features independent of financial statements. Empirical results on a sample of Japanese companies showed that some of the proposed features contain useful information for bankruptcy prediction<sup>16</sup>.

In the domain of financial forecasting, Artificial Neural Networks (ANN)<sup>17</sup> and Support Vector Machines (SVM)<sup>9</sup> have become two of the most widely adopted models, owing to their powerful self-learning capabilities and their effectiveness in capturing complex nonlinear relationships within data. However, both approaches exhibit notable limitations. ANN relies on the backpropagation algorithm, which is prone to becoming trapped in local optima, and its performance is highly sensitive to the tuning of hyperparameters related to network architecture<sup>18</sup>. Although SVM is theoretically well-founded, its computational efficiency is hindered by the iterative optimization of support vectors, resulting in a marked decrease in training speed when applied to large-scale datasets. In contrast, the Extreme Learning Machine (ELM) demonstrates superior overall performance. Its single-hidden-layer feedforward network architecture enables significantly faster training compared to traditional ANN and SVM, while often achieving higher classification accuracy than many advanced classifiers, including SVM<sup>19</sup>. An advanced variant of ELM, the Kernel Extreme Learning Machine (KELM), introduces kernel function techniques to effectively address the precision instability caused by the random assignment of input weights in the original ELM. This enhancement improves the model's ability to extract high-dimensional nonlinear features and significantly boosts its generalization performance<sup>20</sup>. As a result, KELM has been increasingly applied across a wide range of application domains<sup>21</sup>.

Despite the widespread application of Kernel Extreme Learning Machines (KELM) across various domains, extensive research has shown that the performance of KELM classifiers is highly sensitive to two critical parameters: the kernel penalty parameter  $C$  and the kernel bandwidth  $\gamma$ . Therefore, it is essential to appropriately determine these parameters before deploying the classifier in real-world scenarios<sup>22</sup>. Traditionally, grid search has been employed for this purpose; however, this method is prone to local optima and often struggles to define suitable parameter search ranges<sup>21,23</sup>.

To overcome these limitations, a range of optimization techniques has been introduced into the field. Optimization is inherently suited for discovering optimal or near-optimal solutions for unknown variables. For instance, Yu et al.<sup>20</sup> proposed a hybrid forecasting model that combines Complementary Ensemble Empirical Mode Decomposition (CEEMD), Whale Optimization Algorithm (WOA), and KELM for short-term wind power prediction. The WOA was utilized to optimize the two key parameters of KELM, thereby enhancing prediction accuracy and stability. Li et al.<sup>24</sup> applied the Butterfly Optimization Algorithm (BOA) to tune KELM hyperparameters for modeling and analyzing reservoir gas relative permeability. Similarly, Li et al.<sup>25</sup> introduced a Biogeography-Based Optimization (BBO) enhanced KELM approach (BBO-KELM) for ultra-short-term wind power forecasting across different regions. Chen et al.<sup>26</sup> developed an online photovoltaic power prediction model based on chaos analysis, an improved Particle Swarm Optimization (PSO) algorithm, and KELM, aiming to improve forecasting accuracy. Hai et al.<sup>27</sup> leveraged the generalization capability and learning efficiency of KELM, incorporating a multi-strategy enhanced dung beetle optimization (MIDBO) algorithm to refine the model structure for frost heave evaluation in coarse-grained soils—leading to improved convergence accuracy, faster convergence, and enhanced robustness. Zhao et al.<sup>28</sup> proposed a KELM-based regression model for ship motion dynamics, optimized via the Grey Wolf Optimizer (GWO). This approach incorporated a sliding time window for dynamic model updates, effectively combining KELM's fast learning mechanism with GWO's global search capabilities to improve prediction accuracy and timeliness. Li et al.<sup>29</sup> developed a hybrid model based on KELM for predicting the interval time of intermittent pumping wells, where the parameters were optimized using an Improved Brain Storm Optimization algorithm (IBSO-KELM). Han et al.<sup>30</sup> proposed a Short-Term Power Load Forecasting (STPLF) model based on KELM, optimized via an Improved Whale Optimization Algorithm (IWOA), to address the high randomness and low prediction accuracy in electricity load forecasting.

Although existing studies have attempted to apply various optimization algorithms—such as Particle Swarm Optimization (PSO), Grey Wolf Optimizer (GWO), and Whale Optimization Algorithm (WOA)—to parameter tuning of machine learning models to enhance performance in financial risk prediction tasks such as bankruptcy forecasting, these optimization algorithms still commonly suffer from critical shortcomings that are difficult to avoid. As noted in relevant research, "most optimization algorithms remain plagued by issues related to parameter randomness. For example, the coefficient vectors ( $\alpha$ ,  $\beta$ ,  $\delta$ ) in GWO, which control search direction and step size, are randomly initialized. This often leads to dispersed search trajectories in the early iterations, making it difficult to quickly focus on promising regions. In later iterations, due to the overly rapid linear decay of the coefficient vectors, the algorithm prematurely loses its exploration capability, ultimately converging to a local optimum and resulting in premature convergence"<sup>31–33</sup>. These limitations highlight the continued need for more effective and robust optimization strategies. Moreover, the "No Free Lunch" theorem<sup>34</sup> asserts that

no single optimization algorithm performs best across all problems, which further drives research toward the development of improved and problem-specific optimization algorithms.

The Northern goshawk Optimization (NGO) algorithm, inspired by the predatory behavior of northern goshawks, has garnered attention in the field of optimization due to its simplicity and ease of implementation<sup>35</sup>. Existing studies have applied NGO to parameter optimization of models such as Support Vector Machines (SVM)<sup>36–38</sup>. However, its application to Kernel Extreme Learning Machine (KELM) classification problems remains relatively limited. Like many other metaheuristic algorithms, NGO also suffers from common issues such as an imbalance between exploration and exploitation, slow convergence speed, entrapment in local optima, and suboptimal convergence accuracy. To address these shortcomings, this study proposes an enhanced version of NGO—referred to as TIS\_NGO—by incorporating three improvement strategies: (1) a thinking innovation strategy, (2) a prey attack strategy based on differential evolution operations, and (3) a centroid reverse learning boundary control strategy. The objective is to develop an efficient optimization algorithm tailored for KELM parameter tuning, thereby improving the accuracy and stability of bankruptcy prediction models.

Specifically, the thinking innovation strategy enables the algorithm to effectively exploit valuable information within the search space during the exploitation phase, while maintaining stability and population diversity. The prey attack strategy based on differential evolution operations enhances the diversity of search directions, allowing the algorithm to explore broader regions and thus avoid premature convergence to local optima. Meanwhile, the centroid reverse learning boundary control strategy mitigates ineffective searches and accelerates convergence speed. The effectiveness of the proposed TIS\_NGO algorithm is validated through experiments conducted on the CEC2017 and CEC2022 benchmark test suites. In practical bankruptcy prediction tasks, TIS\_NGO is integrated with the KELM model to form a hybrid TIS\_NGO-KELM framework. To ensure the model's generalization capability, ten-fold cross-validation and stratified sampling techniques are employed on the bankruptcy dataset. Experimental results demonstrate that the proposed model significantly outperforms other KELM variants optimized by alternative intelligent algorithms in terms of key metrics such as classification accuracy and Matthews correlation coefficient. The ability of the model to distinguish between bankrupt and non-bankrupt firms is notably enhanced. This deep integration of an improved metaheuristic algorithm with a machine learning model not only overcomes the limitations of traditional parameter tuning approaches, but also offers a novel methodological reference for modeling complex nonlinear problems in the financial domain.

The main contributions of this study are as follows:

- (1) A novel improved algorithm, termed TIS\_NGO, is proposed by integrating three enhancement strategies into the original Northern goshawk Optimization (NGO): the thinking innovation strategy, the prey attack strategy based on differential evolution operations, and the centroid reverse learning boundary control strategy;
- (2) The optimization performance of TIS\_NGO is validated through comprehensive comparisons with other state-of-the-art algorithms on the CEC2017 and CEC2022 benchmark test suites;
- (3) A bankruptcy prediction model, TIS\_NGO-KELM, is developed by combining the proposed TIS\_NGO with Kernel Extreme Learning Machine (KELM). Compared to other KELM-based methods, this model demonstrates superior accuracy in identifying bankrupt firms.

The remainder of this paper is organized as follows: Section "[Northern goshawk optimization and the proposed methodology](#)" introduces the original NGO algorithm and the proposed improvements in TIS\_NGO. Section "[Numerical experiments](#)" presents the validation of TIS\_NGO using the CEC2017 and CEC2022 benchmark suites. Section "[Evaluate the proposed TIS\\_NGO for bankruptcy prediction problem](#)" details the development of the TIS\_NGO-KELM-based bankruptcy prediction model and comparative evaluations. Finally, Section "[Summary and Prospect](#)" concludes the study and discusses future research directions.

## Northern goshawk optimization and the proposed methodology

### Northern goshawk optimization (NGO)

#### Initialization

The Northern goshawk optimization (NGO) algorithm is a population-based metaheuristic method, similar in nature to other heuristic optimization algorithms. NGO begins by randomly generating a population of candidate solutions within the defined search space, as represented by:

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_n \end{bmatrix}_{n \times d} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,j} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,j} & \cdots & x_{2,d} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{i,1} & x_{i,2} & \cdots & x_{i,j} & \cdots & x_{i,d} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,j} & \cdots & x_{n,d} \end{bmatrix}_{n \times d} \quad (1)$$

where  $X$  denotes the population of northern goshawks,  $X_i$  represents the position of the  $i^{th}$  individual,  $X_{i,j}$  denotes the value of the  $j^{th}$  decision variable of the  $i^{th}$  goshawk,  $n$  is the population size, and  $d$  is the dimensionality of the optimization problem.

The initial positions of the goshawks are randomly initialized using the following equation:

$$X_{i,j} = (ub_j - lb_j) \times r_1 + lb_j \quad (2)$$

where  $X_{i,j}$  is the initial value of the  $j^{th}$  decision variable for the  $i^{th}$  candidate solution,  $ub_j$  and  $lb_j$  are the upper and lower bounds, respectively, and  $r_1$  is a uniformly distributed random number in the range  $[0,1]$ .

#### Attack prey (exploration stage)

In the first stage of hunting, each northern goshawk randomly selects a prey and swiftly attacks it. Due to the random selection of prey within the search space, this phase enhances the exploration capability of the NGO algorithm. During this phase, the goshawks perform a global search with the goal of identifying the promising regions of the search space. The corresponding mathematical model is described by Eqs. (3) and (4)<sup>35</sup>:

$$X_{i,j}^{new} = \begin{cases} X_{i,j} + r_2 \times (P_{i,j} - I \times X_{i,j}), & F_{P_i} > F_i \\ X_{i,j} + r_2 \times (X_{i,j} - P_{i,j}), & F_{P_i} \geq F_i \end{cases} \quad (3)$$

$$X_i = \begin{cases} X_i^{new}, & \text{if } F_i^{new} < F_i \\ X_i, & \text{else} \end{cases} \quad (4)$$

where  $X_{i,j}^{new}$  denotes the updated position of the  $i^{th}$  northern goshawk in the  $j^{th}$  dimension during the first phase. The term  $r_2$  is a randomly generated vector of dimension  $1 \times d$  with values uniformly distributed in the range  $[0, 1]$ .  $I$  is a random integer value, either 1 or 2.  $P$  represents the position of the prey,  $F_{P_i}$  and  $F_i$  denote the fitness values of the prey and the goshawk, respectively. If the prey has a better fitness, the goshawk moves toward it; otherwise, it adjusts its position away from the prey. This mechanism promotes global exploration and reduces the risk of premature convergence.

#### Chase prey (exploitation stage)

In the prey pursuit phase, after the northern goshawk attacks the prey, the prey attempts to escape while the goshawk chases it. Due to the fast chasing speed of the goshawk, it can almost always pursue and eventually capture the prey under any circumstances. This behavior simulation enhances the local exploitation capability of the algorithm within the search space. The mathematical model for this phase is described by Eqs. (5) to (7)<sup>35</sup>:

$$X_{i,j}^{new} = X_{i,j} + R \times (2 \times r_3 - 1) \times X_{i,j} \quad (5)$$

$$R = 0.02 \times \left(1 - \frac{t}{T}\right) \quad (6)$$

$$X_i = \begin{cases} X_i^{new}, & \text{if } F_i^{new} < F_i \\ X_i, & \text{else} \end{cases} \quad (7)$$

where  $r_3$  is a vector of dimension  $(1 \times d)$  sampled from a normal distribution,  $t$  denotes the current iteration number, and  $T$  is the maximum number of iterations.  $X_{i,j}^{new}$  represents the updated position of the  $i^{th}$  northern goshawk after the prey pursuit phase, and  $F_i^{new}$  is the corresponding fitness value. The parameter  $R$  decreases linearly over iterations to gradually reduce the step size, thus enabling the algorithm to perform fine-grained local search as it converges.

## Proposed Northern Goshawk optimization

### Thinking innovation strategy (TIS)

In Metaheuristic Search (MHS), the algorithm evaluates performance by optimizing randomly generated positions within the search space and computing their fitness values, which guide subsequent optimization. However, MHS faces several challenges during design and implementation. On one hand, balancing algorithmic performance and computational complexity is critical, requiring a trade-off between optimal performance and solution accuracy. On the other hand, MHS algorithms generally have limited capability to intelligently recognize the current positional demands, lacking mechanisms to avoid redundant calculations. This can lead to multiple fitness evaluations at the same location, thereby increasing computational time<sup>39</sup>. Although the NGO algorithm exhibits strong exploratory capabilities, the complexity of the search space still poses risks of premature convergence to local optima.

To address these issues, this paper introduces a novel Thinking Innovation Strategy (TIS) to enhance NGO. TIS improves the overall performance and individual knowledge of the algorithm by balancing exploration and exploitation through enhanced exploratory capabilities. Specifically, TIS leverages creative thinking to boost the algorithm's ability to identify the demands of the current position, enabling more accurate assessments of the positional value and improving search efficiency. Additionally, TIS periodically updates information of successful individuals to reduce repeated fitness evaluations at the same locations, effectively preventing redundant computations and reducing computational complexity. Moreover, TIS employs random critical thinking to ensure efficient utilization of information within the search range during exploitation, while maintaining optimization stability and diversity. Individuals within TIS primarily rely on their Depth of Knowledge (DOK) to search for solutions<sup>40</sup>. DOK acts as a comprehensive cognitive resource pool supporting critical thinking, and consists of two main layers:

- (1) *DOK1*: Knowledge and experience accumulated through personal growth.
- (2) *DOK2*: Continuously acquired and stored natural phenomena information.

Formally, DOK can be expressed as:  $DOK = DOK1, DOK2$ . This structured representation highlights how DOK enables individuals, through strong cognitive abilities, to effectively solve both traditional simple problems and novel complex challenges. The calculation of DOK is given by Eq. (8):

$$\begin{aligned}
 DOK1 &= C + \left(\frac{FEs}{MaxFEs}\right)^{0.5} \\
 DOK2 &= FEs^{10} \\
 DOK &= DOK1 + DOK2
 \end{aligned}
 \tag{8}$$

where  $DOK1$  represents knowledge and experience accumulated over time,  $DOK2$  denotes continuously acquired information,  $C$  is a constant equal to 0.5,  $FEs$  is the current number of fitness evaluations, and  $MaxFEs = T \times n$  denotes the maximum number of evaluations (with  $T$  as maximum iterations and  $n$  as population size).

Imagination (IM) is defined as the ability to coordinate perceived materials in novel ways to create new mental images. Individual imagination is crucial for innovative thinking and stems from critically evaluating prior knowledge and experience, alongside comprehensive observation and understanding of the subject. Cognitively, it integrates multiple dimensions through intuition, imagery, or combination, unconstrained by existing conclusions, viewpoints, frameworks, or theories, thereby generating new ideas. The computation of IM is expressed in Eq. (9)<sup>40</sup>:

$$IM = \pi \times IE \times rand \tag{9}$$

where  $rand$  is a random number in the interval [0,1], and  $IE$  (Information Event) serves as an important basis for identifying successful individuals. During the initial phase of the algorithm, a successful individual is randomly selected and stored in  $IE$ . As the algorithm progresses, a survival-of-the-fittest mechanism retains outstanding successful individuals, which provide vital guidance for subsequent innovative thinking.

Innovative thinking is rooted in the interplay of imagination, heredity, talent, IQ, and cognition, evolving into a form of original thought emphasizing novelty, divergence, and appropriateness. Critical thinking encourages questioning of current information, fostering new scientific cognition, imagination, and inspiration. When individuals aim to generate novel and valuable outcomes, they often combine or transform their depth of knowledge with rich imagination to innovate. This process effectively guides individuals toward multidirectional thinking, producing diverse outcomes and updated perspectives. The formula for innovative thinking is shown in Eq. (10)<sup>40</sup>:

$$X_{new} = \tan (IM - 0.5 \times \pi) + \left(\frac{X}{DOK} + IE\right) \tag{10}$$

*Prey attack strategy based on differential evolution operation*

In the prey-attacking phase of the Northern Goshawk Optimization (NGO) algorithm, the population update mechanism is relatively simple and fails to fully exploit population diversity and search capabilities. This simplistic update strategy may limit the adaptability of search direction and step size, making the algorithm prone to premature convergence and reducing overall search efficiency<sup>41-43</sup>. To address this limitation, a prey-attacking strategy based on the Differential Evolution (DE) mechanism is proposed, as illustrated in Fig. 1. This

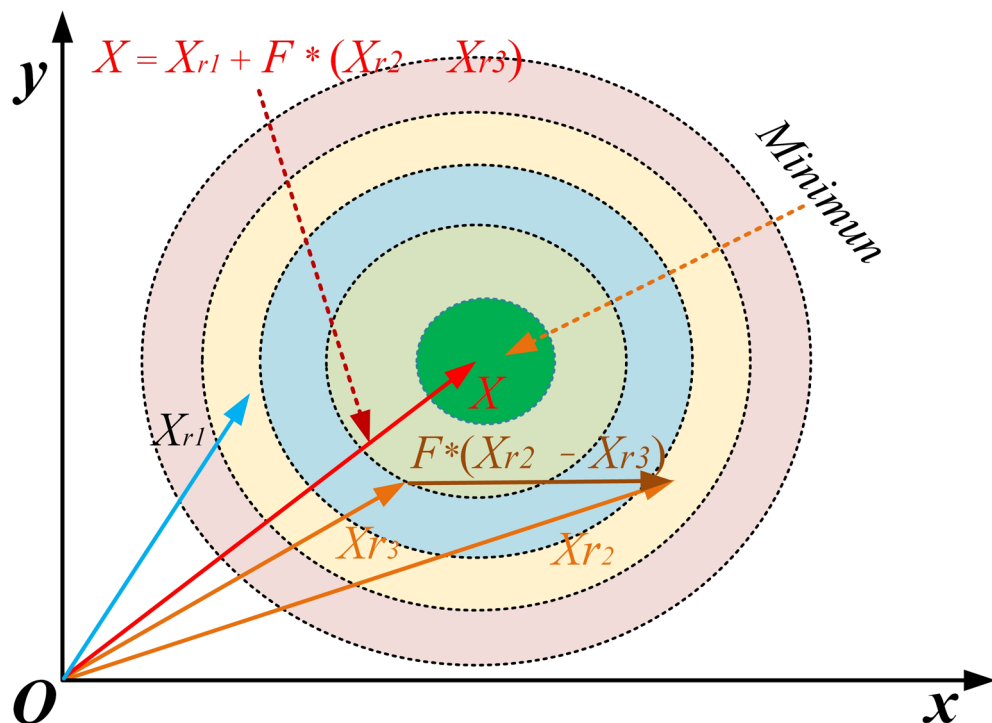


Fig. 1. Illustration of differential evolution operation.

strategy significantly enhances the diversity of search directions by randomly selecting three distinct individuals from the population to perform differential operations. Unlike the traditional update mechanism that relies solely on the current position  $X(t)$  and a fixed parameter, the proposed method utilizes information from different individuals within the population<sup>44–46</sup>. By introducing differential operations, the algorithm is capable of exploring a broader search space and effectively avoiding local optima. The proposed strategy is implemented via Eq. (11):

$$X_{i,j}^{new} = X_{r1} + F \times (X_{r2} - X_{r3}) \tag{11}$$

where  $X_{r1}$ ,  $X_{r2}$  and  $X_{r3}$  are three randomly selected and mutually distinct individuals, and  $F$  denotes the scaling factor (step size) that controls the magnitude of the differential variation. This allows for localized search around individual members and enables the algorithm to dynamically adjust the search step size based on individual differences within the population. As a result, the strategy enhances global exploration ability, increases the likelihood of identifying superior solutions in complex search spaces, and accelerates convergence. This improved mechanism substantially enhances the robustness and efficiency of NGO when tackling complex optimization problems.

*Centroid reverse learning boundary control strategy*

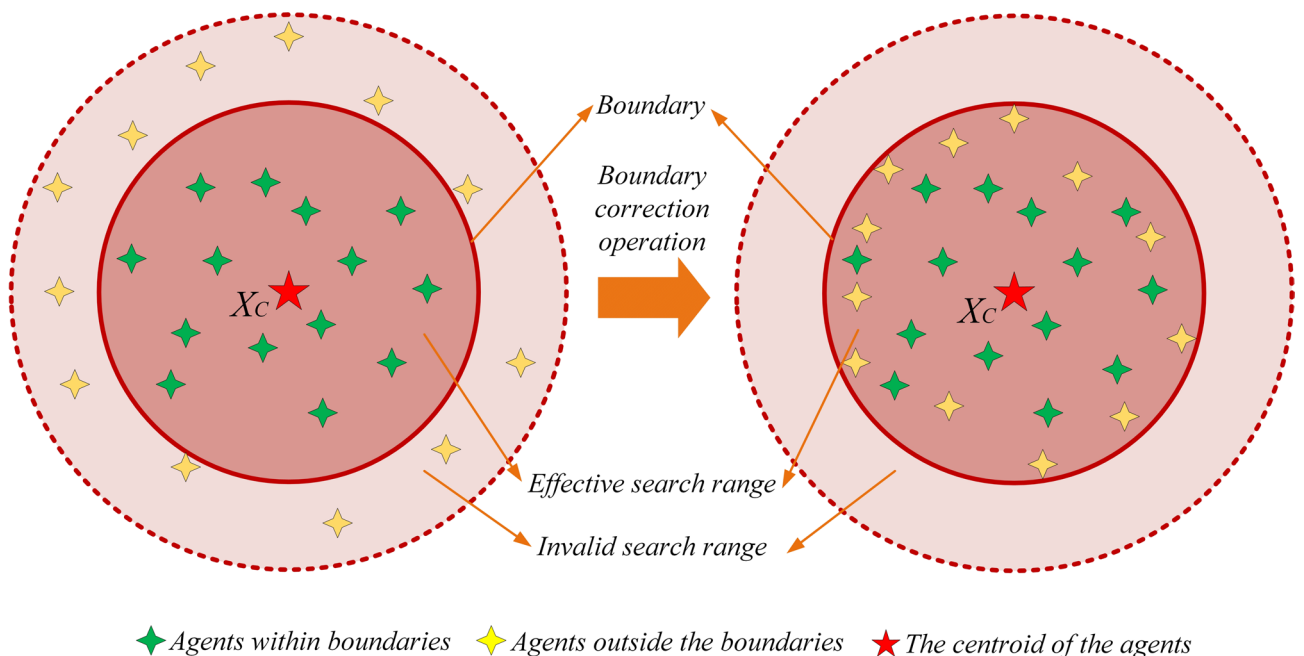
In swarm intelligence algorithms, boundary control plays a critical role by defining the limits of the search space, thereby preventing inefficient searches, premature convergence, and entrapment in local optima<sup>45</sup>. These issues often arise when the search process deviates from feasible solutions or exceeds problem constraints. If the boundaries of the search space are set too narrowly, the algorithm’s ability to thoroughly explore the solution space may be hindered, increasing the risk of falling into local optima<sup>47</sup>. Therefore, the design of an appropriate boundary correction mechanism is essential to enhance the algorithm’s capability to discover the global optimum. Such a mechanism helps prevent ineffective exploration within the search space and reduces time and resource consumption on infeasible solutions. In certain situations, individuals that exceed the boundaries may restrict the search process and limit the utilization of valuable information carried by other members of the population<sup>48</sup>.

To address this issue and maximize the use of positional information from all individuals, a boundary control strategy based on centroid opposition-based learning (COBL) is proposed, as illustrated in Fig. 2. This strategy corrects the positions of out-of-bound individuals by mapping them back into the valid search space, thus avoiding ineffective exploration and accelerating convergence. The correction process is formulated by Eq. (12):

$$X_{i,j}^{new} = \begin{cases} \frac{(lb+X_C)}{2}, & \text{if } X_{i,j}^{new} < lb \\ \frac{(ub+X_C)}{2}, & \text{if } X_{i,j}^{new} > ub \end{cases} \tag{12}$$

where  $X_C$  denotes the centroid of the current population.

In summary, the pseudocode of TIS\_NGO is presented in Algorithm 1.



**Fig. 2.** Schematic diagram of the reverse learning boundary control operation of the centroid.

---

```

1: Initialize Problem Setting (population( $n$ ),  $d$ ,  $ub$ ,  $lb$ ), iterations( $T$ ).
2: Randomly initialize a set of Northern Goshawk' ( $X_i$  ( $i = 1, 2, \dots, n$ )),  $X_{best}$  is the best solution found so far.
3: while  $t = 1:T$  do
4:   Calculate the fitness of the population.
5:   The TIS strategy is introduced to optimize the population in advance (Equations (8)-(10)).
6:   for  $i = 1:n$ 
7:     Exploration stage: Attack prey
8:       Calculate the fitness  $X_{i,j}^{new}$  using Equation (3).
9:       Update the position of the current individual using Equation (4) and Equation (11).
10:      Using equation (12) for boundary adjustment
11:     Exploitation: Chase prey
12:       Calculate the fitness  $x_{i,j}^{new2}$  using Equations (5)-(6).
13:       Update the position of the current individual using Equation (7).
14:       Using equation (12) for boundary adjustment
15:     End for
14:   Update the best solution found so far  $X_{best}$ .
15: End while
16: Return  $X_{best}$ .

```

---

**Algorithm 1:** Pseudo-Code of TIS\_NGOæ.

*Computational time complexity*

An algorithm's efficacy is vital; assessing its temporal complexity holds equal significance. For numerous optimization duties, a method must show superb results while maintaining exceptional immediacy. Temporal complexity illustrates runtime growth relative to input magnitude. Examining an optimizer's time cost aids in approximating computational expense for vast-scale issues. Within the standard NGO, defined control parameters demand  $O(N \times dim)$  operations, with  $N$  being group quantity and  $dim$  indicating dimensionality. Initially, the procedure needs  $O(N \times dim)$  duration. Across  $T$  cycles, refreshing member locations consumes  $O(T \times N \times dim)$ . Thus, NGO's complete computational burden is  $O(T \times N \times dim)$ . In the devised TIS\_NGO, solely the location renewal tactic and goal function assessment approach are enhanced, adding no further complexity elements, so time expenditure stays  $O(T \times N \times dim)$ .

## Numerical experiments

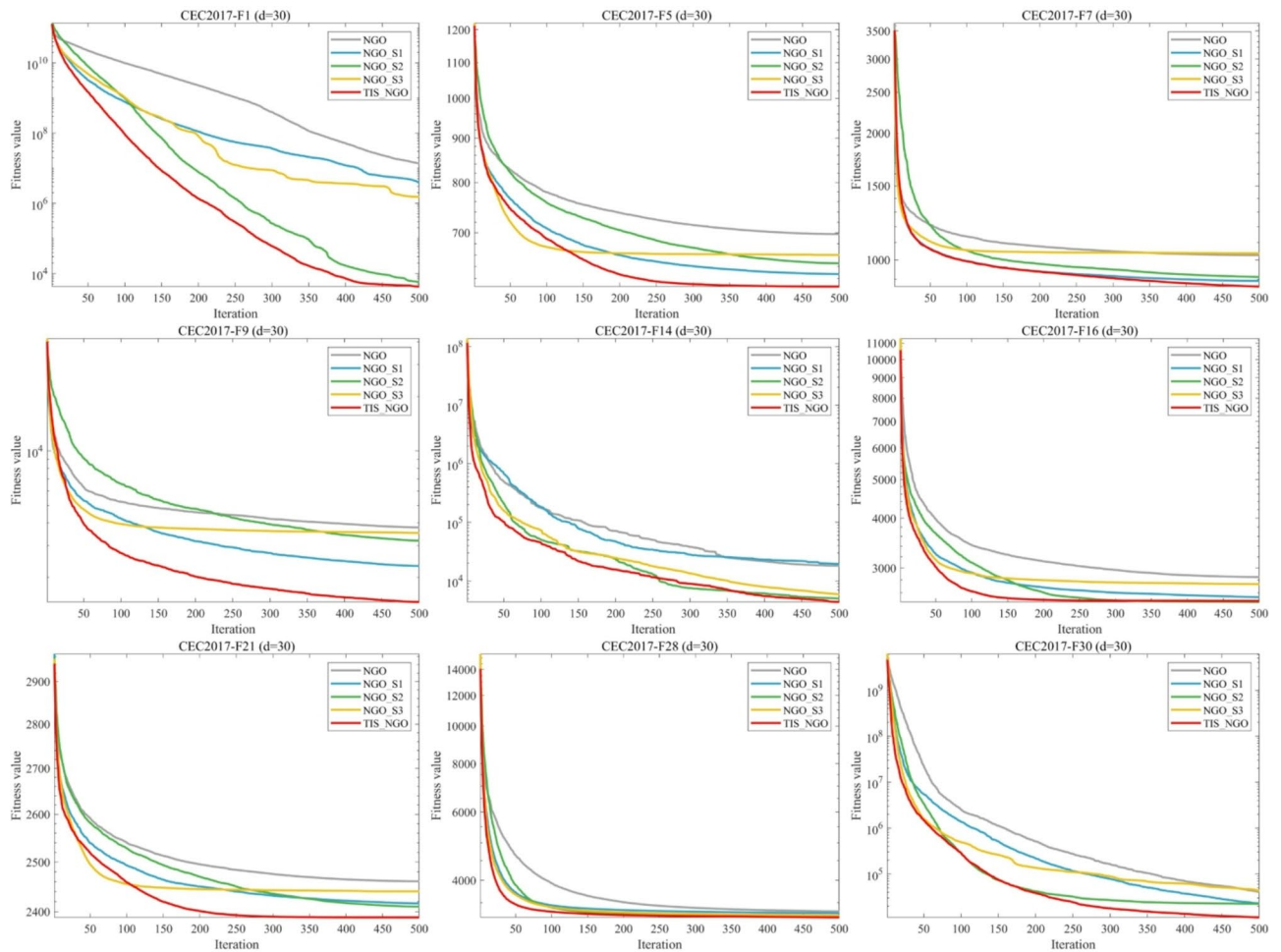
In this subsection, the effectiveness of the proposed TIS\_NGO is validated through comparative experiments with other algorithms on the CEC2017 benchmark functions<sup>49</sup>.

### Ablation experiment

To evaluate the individual contributions and synergistic effects of three enhancement strategies—Thinking Innovation Strategy (S1), Prey Attack Strategy Based on Differential Evolution Operation (S2), and Centroid Reverse Learning Boundary Control Strategy (S3)—on the baseline NGO algorithm, this section conducts ablation experiments using the CEC 2017 test benchmark (with dimension  $d=30$ ). Four comparative variants are designed: NGO\_S1 (with S1 only), NGO\_S2 (with S2 only), NGO\_S3 (with S3 only), and TIS\_NGO, which integrates all three strategies. The experimental results are presented in Figs. 3 and 4.

As observed from the convergence curves in Fig. 3, on typical test functions such as CEC2017-F1, F5, and F7, all single-strategy enhanced variants (NGO\_S1, NGO\_S2, NGO\_S3) outperform the original NGO in convergence performance. Among them, NGO\_S2 (with S2 only) demonstrates stronger global exploration capability in the early iterations, converging faster than NGO\_S1 (S1 only) and NGO\_S3 (S3 only). In contrast, TIS\_NGO, through the synergy of the three strategies, not only rapidly approaches the optimal solution region in the early stages but also maintains stable local search accuracy in later phases. Its convergence curve remains the lowest throughout, significantly outperforming both the single-strategy variants and the original NGO, validating that the multi-strategy integration enhances the algorithm's balance between exploration and exploitation.

The average ranking results in Fig. 4 further quantify this advantage: the original NGO has an average ranking of 3.93. Among the three single-strategy variants, NGO\_S2 performs the best (average ranking 3.70), followed by NGO\_S1 (3.13) and NGO\_S3 (2.73). In comparison, TIS\_NGO achieves a markedly lower average ranking of only 1.50, far surpassing all other variants. This indicates a significant synergistic effect among the three



**Fig. 3.** Convergence curves of NGO with different strategy improvements.

strategies: S1 improves information utilization efficiency through Depth of Knowledge (DOK) and Imagination (IM), S2 enhances search direction diversity via differential evolution, and S3 reduces ineffective search through boundary control. Their combination effectively addresses issues such as the slow convergence and tendency to fall into local optima observed in the original NGO, thereby establishing a solid foundation for the superior performance of TIS\_NGO.

### Algorithm parameter settings

This section evaluates the performance of the proposed TIS\_NGO using the most challenging benchmark suite from the CEC2017 numerical optimization competition, and compares it with several state-of-the-art algorithms. The comparison algorithms include Particle Swarm Optimization (PSO)<sup>50</sup>, Differential Evolution (DE)<sup>51</sup>, Grey Wolf Optimizer (GWO)<sup>52</sup>, Whale Optimization Algorithm (WOA)<sup>53</sup>, Dung Beetle Optimizer (DBO)<sup>54</sup>, Black-winged Kite Algorithm (BKA)<sup>55</sup>, Animated Oat Optimization (AOO)<sup>56</sup>, and Northern Goshawk Optimization (NGO)<sup>35</sup>. The specific parameter settings for each algorithm are listed in Table 1. To ensure fair comparison and eliminate randomness, the population size was uniformly set to 30, and the maximum number of iterations was fixed at 500 for all algorithms. Each algorithm was independently executed 30 times. The performance metrics, including the average value (Ave), standard deviation (Std), and ranking (Rank), were recorded, with the best results highlighted in bold. All experiments were conducted on a computer running Windows 10 with a 13th Gen Intel(R) Core(TM) i5-13,400 CPU @ 2.5 GHz, 16 GB of RAM, using MATLAB R2022b as the software platform.

### Experimental results and analysis of CEC2017 and CEC2022 test suite

The performance of TIS\_NGO is evaluated in this section using two benchmark suites: CEC2017 and CEC2022. To comprehensively assess its effectiveness, experiments were conducted on the CEC2017 benchmark with 30 dimensions ( $d=30$ ) and the CEC2022 benchmark with 20 dimensions ( $d=20$ ). The experimental results are summarized in Tables 2 and 3, which report the average value (Ave), standard deviation (Std), and the ranking (Rank) for each algorithm. The best results are highlighted in bold. The convergence curves are illustrated in Fig. 5, demonstrating the convergence speed of each algorithm.

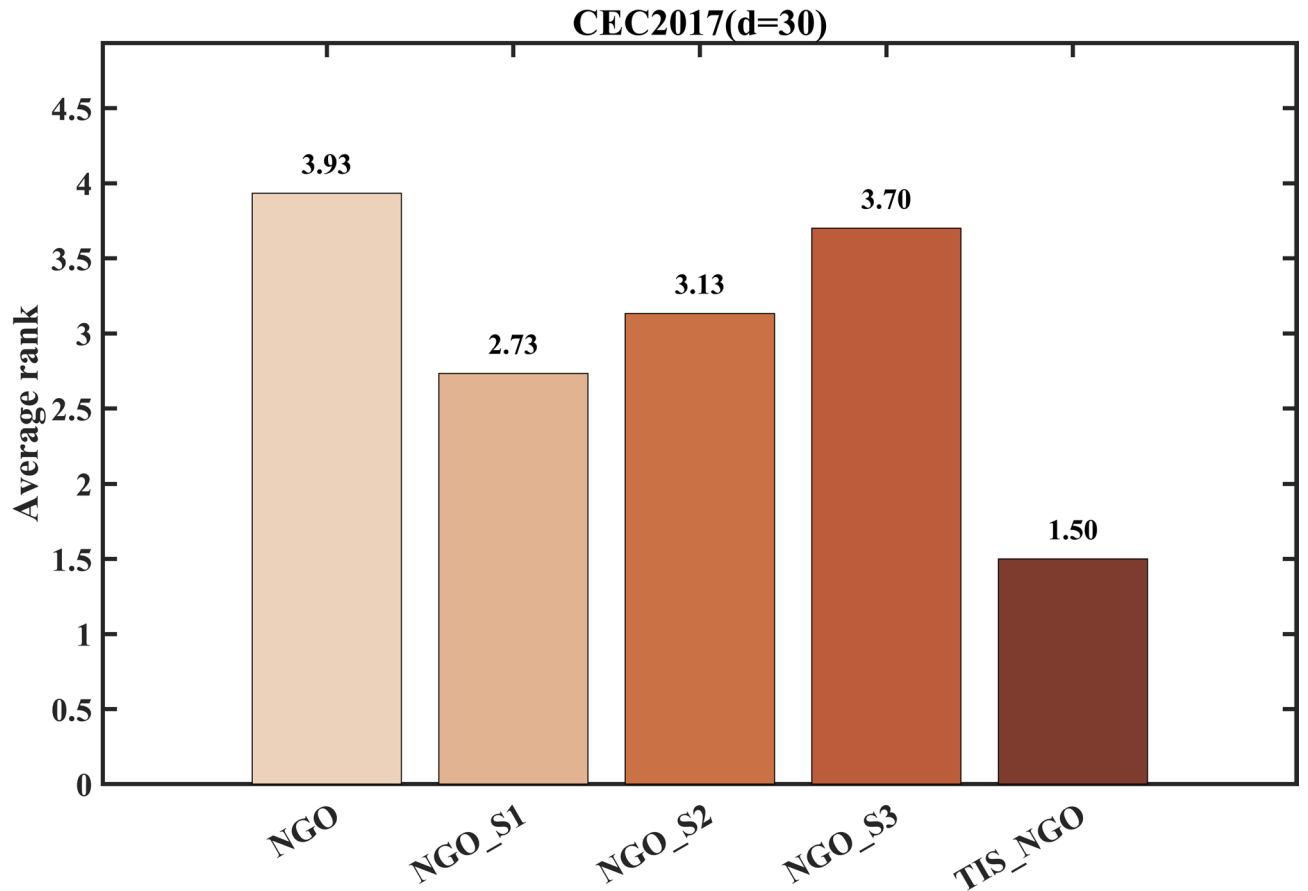


Fig. 4. Average rankings of different NGO variant algorithms on test functions.

Algorithms	Name of the parameter	Value of the parameter
PSO	$c_1, c_2, w$	1.5, 1.5, 0.8
DE	$F, Cr$	0.8, 0.1
GWO	$a$	[0,2]
WOA	$r, l, a$	[0,1], [-1,1], Linear reduction from 2 to 1
DBO	$P_{percent}$	0.2
BKA	$P, r$	0.9, [0,1]
AOO	$K, \theta$	[0.5,1.5], [0, $\pi$ ]
NGO	$I, R$	{1, 2}, [0,0.02]
TIS_NGO	$I, R, F$	{1, 2}, [0,0.02], [0,0.5]

Table 1. Compare algorithm parameter settings.

As illustrated in the tables, TIS\_NGO demonstrates superior performance in terms of both average fitness values and standard deviations across the CEC2017 and CEC2022 benchmark sets, outperforming the standard NGO and other comparative algorithms. Specifically, TIS\_NGO achieves the best results on 23 out of 30 functions in the CEC2017 test suite, with only 7 functions (F7, F15, F17, F19, F22, F26, and F27) not being the best. In contrast, the standard NGO yields the best performance on only 5 functions (F15, F17, F19, F22, and F27). On the CEC2022 benchmark suite, which includes 12 functions, TIS\_NGO obtains the best average fitness values on 7 functions (F2–F5, F8, F9, and F11), while the standard NGO achieves the best results on only F6, F10, and F12. In addition, TIS\_NGO consistently achieves the lowest standard deviations (highlighted in bold), further demonstrating its superior robustness and stability over the standard NGO and other peer algorithms.

As shown in the average convergence curves (Fig. 5), the average fitness values of all algorithms tend to decrease as the number of iterations increases, indicating their ability to converge toward better solutions. TIS\_NGO exhibits exceptional performance on various types of optimization problems, including unimodal, multimodal, hybrid, and composite functions.

F~	Metric	PSO	DE	GWO	WOA	DBO	BKA	AOO	NGO	TIS_NGO
F1	Ave	1.5273E+09	2.7672E+07	2.8848E+09	5.7844E+09	2.5742E+08	1.2045E+10	9.1247E+04	2.2144E+07	<b>5.6759E+03</b>
	Std	1.3656E+09	5.1295E+06	1.8016E+09	1.8301E+09	2.0045E+08	1.3584E+10	5.0152E+04	5.3911E+07	<b>5.5389E+03</b>
	Rank	6	4	7	8	5	9	2	3	<b>1</b>
F2	Ave	1.0934E+31	9.2474E+27	1.0254E+33	4.5532E+36	5.9732E+32	5.1396E+41	1.3813E+17	1.4806E+21	<b>2.3631E+14</b>
	Std	3.5203E+31	1.3332E+28	3.5567E+33	2.2825E+37	1.5736E+33	2.8151E+42	6.8247E+17	8.1034E+21	<b>7.5898E+14</b>
	Rank	4	6	7	9	8	5	2	3	<b>1</b>
F3	Ave	5.9739E+04	1.2958E+05	6.3345E+04	2.3876E+05	8.9944E+04	3.5558E+04	2.4521E+04	6.2719E+04	<b>1.5994E+04</b>
	Std	1.6274E+04	2.0718E+04	1.1572E+04	6.0949E+04	2.0656E+04	1.9482E+04	8.6152E+03	8.0849E+03	<b>5.1716E+03</b>
	Rank	4	8	5	9	7	3	2	6	<b>1</b>
F4	Ave	6.3360E+02	6.3876E+02	6.1327E+02	1.3264E+03	7.2737E+02	2.5265E+03	5.0057E+02	5.2532E+02	<b>4.7586E+02</b>
	Std	1.0579E+02	2.7517E+01	7.8160E+01	3.4563E+02	2.0454E+02	3.7099E+03	1.8433E+01	<b>1.8224E+01</b>	3.8905E+01
	Rank	5	6	4	9	7	8	2	3	<b>1</b>
F5	Ave	7.1899E+02	6.8268E+02	6.3604E+02	8.6641E+02	7.4829E+02	7.4414E+02	6.2561E+02	6.9744E+02	<b>6.0179E+02</b>
	Std	2.5706E+01	<b>1.4609E+01</b>	4.4741E+01	7.1220E+01	6.6016E+01	5.2019E+01	2.5707E+01	3.0848E+01	2.7205E+01
	Rank	6	4	3	9	8	7	2	5	<b>1</b>
F6	Ave	6.2319E+02	6.0385E+02	6.1195E+02	6.8527E+02	6.5052E+02	6.6045E+02	6.3087E+02	6.3893E+02	<b>6.0184E+02</b>
	Std	8.9557E+00	<b>6.1553E-01</b>	5.1746E+00	1.1910E+01	1.1618E+01	8.8657E+00	1.0926E+01	1.2059E+01	1.4676E+00
	Rank	4	2	3	9	7	8	5	6	<b>1</b>
F7	Ave	1.0020E+03	9.4546E+02	9.0059E+02	1.3382E+03	1.0373E+03	1.2270E+03	<b>8.6645E+02</b>	1.0315E+03	8.7078E+02
	Std	3.2149E+01	<b>1.2334E+01</b>	3.9340E+01	6.5962E+01	8.6066E+01	6.7450E+01	3.3179E+01	4.8416E+01	4.5795E+01
	Rank	5	4	3	9	6	8	<b>1</b>	7	2
F8	Ave	1.0121E+03	9.9148E+02	9.1027E+02	1.0711E+03	1.0211E+03	9.7395E+02	9.1524E+02	9.6627E+02	<b>8.9817E+02</b>
	Std	3.0888E+01	<b>1.4473E+01</b>	3.2697E+01	4.6821E+01	5.7743E+01	2.0727E+01	3.6158E+01	2.1837E+01	2.9900E+01
	Rank	8	6	2	9	7	5	3	4	<b>1</b>
F9	Ave	1.9126E+03	3.8167E+03	2.4679E+03	1.1970E+04	6.7373E+03	5.7805E+03	3.5270E+03	3.7353E+03	<b>1.7255E+03</b>
	Std	7.3887E+02	6.2997E+02	7.5545E+02	3.0127E+03	1.7228E+03	1.5170E+03	1.1891E+03	<b>6.1033E+02</b>	7.0135E+02
	Rank	2	5	3	9	8	7	4	6	<b>1</b>
F10	Ave	7.6060E+03	6.8597E+03	5.2148E+03	7.5397E+03	6.4048E+03	5.8044E+03	4.8242E+03	5.4371E+03	<b>4.0180E+03</b>
	Std	4.6006E+02	<b>3.3214E+02</b>	1.5953E+03	8.5708E+02	1.3067E+03	1.0068E+03	6.1942E+02	4.3850E+02	4.4707E+02
	Rank	9	7	3	8	6	5	2	4	<b>1</b>
F11	Ave	1.5147E+03	2.3015E+03	2.8659E+03	9.1030E+03	2.3741E+03	2.1627E+03	1.2896E+03	1.2566E+03	<b>1.2020E+03</b>
	Std	1.0847E+02	4.4864E+02	1.2609E+03	3.7461E+03	1.3009E+03	2.7795E+03	4.3459E+01	<b>3.5437E+01</b>	4.3357E+01
	Rank	5	7	8	9	6	4	3	2	<b>1</b>
F12	Ave	1.4727E+08	3.3115E+07	1.1855E+08	5.3929E+08	3.9260E+07	9.8945E+08	1.4208E+07	1.6580E+06	<b>2.4164E+05</b>
	Std	4.3282E+08	8.5564E+06	1.0252E+08	3.4249E+08	4.1888E+07	2.0777E+09	1.7432E+07	1.1820E+06	<b>2.4806E+05</b>
	Rank	8	5	7	9	4	6	3	2	<b>1</b>
F13	Ave	9.0488E+06	6.3076E+06	8.3784E+06	1.1561E+07	3.8507E+07	7.9214E+07	9.7408E+04	1.7558E+04	<b>1.5050E+04</b>
	Std	1.7538E+07	3.2801E+06	2.4339E+07	1.3935E+07	1.4384E+08	3.0404E+08	7.0215E+04	<b>1.4284E+04</b>	1.9805E+04
	Rank	7	8	5	9	6	4	3	2	<b>1</b>
F14	Ave	1.1666E+05	4.9376E+05	7.0611E+05	2.7894E+06	5.2301E+05	3.5486E+04	4.5441E+04	3.5629E+04	<b>5.1684E+03</b>
	Std	1.0297E+05	3.3799E+05	7.5489E+05	2.1458E+06	7.0599E+05	1.4611E+05	4.2705E+04	3.1934E+04	<b>3.6646E+03</b>
	Rank	5	8	7	9	6	2	3	4	<b>1</b>
F15	Ave	2.2592E+05	1.4456E+06	3.3998E+06	5.8425E+06	7.4183E+05	3.8098E+04	4.6812E+04	<b>4.9913E+03</b>	1.0595E+04
	Std	2.7530E+05	7.8343E+05	1.3675E+07	6.5136E+06	3.6901E+06	2.3384E+04	3.8984E+04	<b>3.7547E+03</b>	8.8158E+03
	Rank	6	8	7	9	5	3	4	<b>1</b>	2
F16	Ave	3.0388E+03	3.0051E+03	2.6851E+03	4.3253E+03	3.3465E+03	3.2166E+03	2.7270E+03	2.8299E+03	<b>2.4099E+03</b>
	Std	3.2548E+02	<b>1.5651E+02</b>	3.0418E+02	6.3632E+02	4.3439E+02	5.1376E+02	2.6735E+02	2.2374E+02	2.6334E+02
	Rank	5	6	2	9	8	7	3	4	<b>1</b>
F17	Ave	2.1077E+03	2.2046E+03	2.0781E+03	2.7450E+03	2.6123E+03	2.4217E+03	2.1531E+03	<b>2.0429E+03</b>	2.0586E+03
	Std	1.6572E+02	1.0311E+02	1.9112E+02	2.4782E+02	2.4424E+02	2.2960E+02	2.1961E+02	<b>9.8021E+01</b>	1.8458E+02
	Rank	3	6	4	9	8	7	5	<b>1</b>	2
F18	Ave	2.1554E+06	1.5684E+06	1.5065E+06	1.4424E+07	3.0030E+06	1.5822E+05	1.0323E+06	2.0792E+05	<b>1.4478E+05</b>
	Std	2.6540E+06	8.0950E+05	1.3859E+06	1.6373E+07	4.2068E+06	1.7849E+05	1.0915E+06	1.5653E+05	<b>1.0106E+05</b>
	Rank	7	8	5	9	6	<b>1</b>	4	3	2

Continued

F~	Metric	PSO	DE	GWO	WOA	DBO	BKA	AOO	NGO	TIS_NGO
F19	Ave	4.8941E+05	1.1991E+06	2.4604E+06	2.9056E+07	4.5130E+06	5.5082E+06	5.6791E+05	<b>6.9051E+03</b>	9.5900E+03
	Std	4.3619E+05	7.4968E+05	7.3729E+06	2.9259E+07	5.0861E+06	1.8896E+07	6.1275E+05	<b>4.2159E+03</b>	8.1669E+03
	Rank	4	7	6	9	8	3	5	<b>1</b>	2
F20	Ave	2.5249E+03	2.4991E+03	2.4600E+03	2.9022E+03	2.7384E+03	2.5391E+03	2.5241E+03	2.4730E+03	<b>2.4489E+03</b>
	Std	1.5004E+02	1.0925E+02	1.7007E+02	1.5262E+02	1.7889E+02	1.9822E+02	1.7348E+02	<b>8.1003E+01</b>	1.5446E+02
	Rank	4	5	<b>1</b>	9	8	6	7	3	2
F21	Ave	2.5017E+03	2.4881E+03	2.4101E+03	2.6437E+03	2.5637E+03	2.5369E+03	2.4130E+03	2.4569E+03	<b>2.3905E+03</b>
	Std	2.3316E+01	<b>1.4524E+01</b>	3.0522E+01	5.7454E+01	5.7034E+01	5.6139E+01	2.7344E+01	2.6146E+01	2.5049E+01
	Rank	6	5	2	9	8	7	3	4	<b>1</b>
F22	Ave	5.3304E+03	3.7408E+03	5.3292E+03	8.4695E+03	4.3575E+03	6.5811E+03	4.5535E+03	<b>2.3379E+03</b>	4.1402E+03
	Std	3.0692E+03	8.3318E+02	1.7618E+03	1.3408E+03	2.0400E+03	1.4376E+03	2.0394E+03	<b>3.1969E+01</b>	1.7927E+03
	Rank	7	4	6	9	5	8	3	<b>1</b>	2
F23	Ave	2.9396E+03	2.8366E+03	2.7928E+03	3.1540E+03	3.0358E+03	3.1289E+03	2.7970E+03	2.8214E+03	<b>2.7606E+03</b>
	Std	8.6149E+01	<b>1.5100E+01</b>	4.1481E+01	8.5081E+01	7.3061E+01	1.1539E+02	4.8136E+01	4.2620E+01	3.3127E+01
	Rank	6	5	3	9	7	8	2	4	<b>1</b>
F24	Ave	3.1394E+03	3.0486E+03	2.9594E+03	3.2904E+03	3.1737E+03	3.2678E+03	2.9705E+03	2.9808E+03	<b>2.9234E+03</b>
	Std	7.5022E+01	<b>2.3884E+01</b>	5.6373E+01	1.1168E+02	7.8325E+01	1.2465E+02	4.3298E+01	3.7507E+01	3.0061E+01
	Rank	6	5	2	9	7	8	3	4	<b>1</b>
F25	Ave	2.9599E+03	2.9831E+03	3.0251E+03	3.2235E+03	2.9829E+03	3.3983E+03	2.9125E+03	2.9500E+03	<b>2.8966E+03</b>
	Std	2.9244E+01	1.8383E+01	7.8820E+01	7.7385E+01	6.0365E+01	7.0387E+02	2.0137E+01	2.8133E+01	<b>1.7772E+01</b>
	Rank	4	6	7	9	5	8	2	3	<b>1</b>
F26	Ave	5.1412E+03	5.3308E+03	4.9515E+03	8.3932E+03	6.6226E+03	7.5436E+03	4.6209E+03	<b>3.9598E+03</b>	4.7525E+03
	Std	1.0034E+03	<b>4.3360E+02</b>	4.5144E+02	1.1633E+03	9.7568E+02	2.0415E+03	1.0486E+03	1.2256E+03	8.1010E+02
	Rank	5	6	4	9	7	8	3	<b>1</b>	2
F27	Ave	3.2779E+03	3.2738E+03	3.2710E+03	3.4763E+03	3.3053E+03	3.4300E+03	3.2501E+03	<b>3.2334E+03</b>	3.2349E+03
	Std	4.0002E+01	<b>8.6616E+00</b>	3.0424E+01	1.6030E+02	4.2900E+01	1.5899E+02	1.8340E+01	1.6890E+01	1.2152E+01
	Rank	4	6	5	9	7	8	3	<b>1</b>	2
F28	Ave	3.4073E+03	3.3904E+03	3.4851E+03	3.8814E+03	3.5747E+03	3.8017E+03	3.2580E+03	3.3121E+03	<b>3.2178E+03</b>
	Std	1.3690E+02	2.5957E+01	9.7476E+01	2.1596E+02	5.8269E+02	6.0254E+02	3.6695E+01	<b>2.4437E+01</b>	2.9872E+01
	Rank	4	5	7	9	6	8	2	3	<b>1</b>
F29	Ave	4.0894E+03	4.0426E+03	3.9905E+03	5.2009E+03	4.5493E+03	4.7632E+03	4.0518E+03	4.1597E+03	<b>3.7444E+03</b>
	Std	2.5817E+02	<b>1.0910E+02</b>	2.6447E+02	5.5949E+02	3.9987E+02	5.0323E+02	2.6302E+02	1.8446E+02	2.0386E+02
	Rank	3	5	2	9	7	8	4	6	<b>1</b>
F30	Ave	2.5662E+06	1.5117E+06	1.0228E+07	7.9807E+07	4.3734E+06	5.4525E+07	3.3387E+06	3.8541E+04	<b>1.1432E+04</b>
	Std	2.0586E+06	7.3731E+05	8.3189E+06	7.8863E+07	5.8792E+06	1.3731E+08	1.6311E+06	5.0301E+04	<b>4.1196E+03</b>
	Rank	4	3	8	9	5	6	7	2	<b>1</b>
Ave,Rank		5.20	5.67	4.60	8.93	6.60	6.17	3.23	3.30	<b>1.30</b>
Friedman		5	6	4	9	8	7	2	3	<b>1</b>

**Table 2.** Experimental results of CEC2017 test suite.

Specifically, on unimodal functions such as CEC2017-F1, TIS\_NGO demonstrates a significantly faster convergence rate in the early stages compared to the baseline algorithms, and it continues to improve solution quality in later iterations, reflecting its high search efficiency. In contrast, standard NGO, GWO, WOA, and PSO exhibit slower convergence behavior. For multimodal functions with numerous local optima (e.g., CEC2017-F9 and CEC2022-F5), TIS\_NGO achieves rapid global search, effectively approaching the optimal region in early iterations with large steps, while avoiding premature convergence—a common issue in PSO and WOA.

In the case of hybrid and composite functions, such as CEC2017-F30, TIS\_NGO also demonstrates outstanding global exploration and local exploitation capabilities. While comparison algorithms often suffer from stagnation or getting trapped in local optima, TIS\_NGO, empowered by its three synergistic enhancement strategies, dynamically adjusts its search behavior throughout the optimization process, enabling continuous progress toward global optima. This advantage allows TIS\_NGO to maintain high optimization efficiency, even when tackling complex, multimodal, or high-dimensional problems.

Furthermore, to better visualize the performance of each algorithm across the test functions, radar charts were employed to illustrate the ranking distribution of each algorithm on individual test functions, as shown in Fig. 6. Additionally, the average rankings of all algorithms on the CEC2017 and CEC2022 benchmark suites were computed and are presented in Fig. 7.

As shown in Fig. 6, the proposed TIS\_NGO algorithm exhibits the smallest enclosed area on the radar charts for both the CEC2017 and CEC2022 test functions, with rankings predominantly concentrated within

F~	Metric	PSO	DE	GWO	WOA	DBO	BKA	AOO	NGO	TIS_NGO
F1	Ave	6.4907E+03	3.4172E+04	1.6332E+04	3.5604E+04	3.3879E+04	7.0048E+03	<b>4.1505E+02</b>	1.4153E+04	5.4133E+02
	Std	3.2855E+03	6.8988E+03	5.8791E+03	1.0921E+04	1.0572E+04	8.5040E+03	<b>1.2601E+02</b>	3.4638E+03	1.6424E+02
	Rank	4	9	6	8	7	3	<b>1</b>	5	2
F2	Ave	4.7907E+02	4.7845E+02	5.0350E+02	6.4324E+02	4.9554E+02	5.7147E+02	4.5344E+02	4.6844E+02	<b>4.4361E+02</b>
	Std	3.4482E+01	<b>1.0378E+01</b>	4.7765E+01	5.9444E+01	4.4294E+01	2.5907E+02	1.7120E+01	1.8541E+01	2.4234E+01
	Rank	4	5	7	9	6	8	2	3	<b>1</b>
F3	Ave	6.1338E+02	6.0069E+02	6.0534E+02	6.7175E+02	6.3502E+02	6.5429E+02	6.1862E+02	6.1979E+02	<b>6.0021E+02</b>
	Std	4.8626E+00	<b>1.6714E-01</b>	2.6745E+00	1.1636E+01	1.3189E+01	6.8787E+00	1.0017E+01	1.0194E+01	3.0382E-01
	Rank	4	2	3	9	7	8	5	6	<b>1</b>
F4	Ave	9.0905E+02	9.2724E+02	8.6596E+02	9.4371E+02	9.0837E+02	8.8106E+02	8.5960E+02	8.6798E+02	<b>8.5820E+02</b>
	Std	1.5344E+01	1.2293E+01	3.2746E+01	3.2002E+01	3.4193E+01	2.3081E+01	2.0127E+01	<b>1.2228E+01</b>	1.6699E+01
	Rank	7	8	3	9	6	5	2	4	<b>1</b>
F5	Ave	1.0179E+03	2.1963E+03	1.2383E+03	4.3427E+03	2.1290E+03	2.2264E+03	1.6184E+03	1.5564E+03	<b>9.5359E+02</b>
	Std	<b>5.4347E+01</b>	3.9724E+02	2.9724E+02	1.7377E+03	5.9363E+02	3.6002E+02	8.0172E+02	2.5022E+02	6.1454E+01
	Rank	2	7	3	9	6	8	4	5	<b>1</b>
F6	Ave	2.2191E+06	4.0002E+06	3.7599E+06	9.2020E+06	3.9039E+05	6.0293E+06	6.8240E+03	<b>4.1285E+03</b>	5.2937E+03
	Std	1.7081E+06	1.7530E+06	1.1346E+07	1.4489E+07	1.2572E+06	2.6762E+07	5.4564E+03	<b>1.9631E+03</b>	4.0148E+03
	Rank	7	9	6	8	5	4	3	<b>1</b>	2
F7	Ave	2.1283E+03	<b>2.0621E+03</b>	2.1054E+03	2.2336E+03	2.1336E+03	2.1415E+03	2.0922E+03	2.0913E+03	2.0653E+03
	Std	5.9782E+01	<b>1.0391E+01</b>	4.2284E+01	5.1740E+01	4.4943E+01	4.1969E+01	4.5675E+01	1.5535E+01	4.6567E+01
	Rank	6	<b>1</b>	5	9	7	8	3	4	2
F8	Ave	2.2892E+03	2.2289E+03	2.2521E+03	2.2988E+03	2.2996E+03	2.2807E+03	2.2517E+03	2.2359E+03	<b>2.2252E+03</b>
	Std	6.0407E+01	<b>1.2604E+00</b>	4.4092E+01	7.4906E+01	8.7415E+01	6.5925E+01	4.0503E+01	2.2226E+01	5.7480E+00
	Rank	8	2	4	9	7	6	5	3	<b>1</b>
F9	Ave	2.4948E+03	2.4821E+03	2.5145E+03	2.6086E+03	2.5147E+03	2.5585E+03	2.4832E+03	2.4808E+03	<b>2.4808E+03</b>
	Std	1.8332E+01	5.1663E-01	1.9728E+01	5.8482E+01	3.8563E+01	1.3897E+02	2.5974E+00	1.8988E-02	<b>8.1330E-04</b>
	Rank	5	3	8	9	6	7	4	2	<b>1</b>
F10	Ave	3.9831E+03	2.5461E+03	3.8391E+03	5.1651E+03	3.0379E+03	4.1893E+03	3.5569E+03	<b>2.5433E+03</b>	2.7494E+03
	Std	9.9236E+02	8.8954E+01	1.0226E+03	1.1151E+03	9.7344E+02	9.9842E+02	8.2573E+02	<b>7.1887E+01</b>	2.5330E+02
	Rank	7	3	6	9	4	8	5	<b>1</b>	2
F11	Ave	3.3656E+03	3.1084E+03	3.7011E+03	3.8114E+03	3.1388E+03	4.1225E+03	2.9463E+03	2.9338E+03	<b>2.9271E+03</b>
	Std	2.8373E+02	1.4029E+02	4.8373E+02	3.1151E+02	1.7726E+02	1.3546E+03	2.5028E+02	<b>5.2251E+01</b>	1.6033E+02
	Rank	6	4	7	9	5	8	2	3	<b>1</b>
F12	Ave	2.9879E+03	2.9757E+03	2.9833E+03	3.1240E+03	3.0277E+03	3.0743E+03	2.9701E+03	<b>2.9522E+03</b>	2.9579E+03
	Std	3.7498E+01	<b>5.3385E+00</b>	2.6283E+01	1.3368E+02	5.4739E+01	6.9236E+01	1.9297E+01	1.2649E+01	1.6385E+01
	Rank	6	4	5	9	7	8	3	<b>1</b>	2
Ave,Rank		5.50	4.75	5.25	8.83	6.08	6.75	3.25	3.17	<b>1.42</b>
Friedman		6	4	5	9	7	8	3	2	<b>1</b>

**Table 3.** Experimental results of CEC2022 test suite.

the top two positions and no worst-case rankings. This clearly demonstrates the effectiveness of the TIS\_NGO algorithm, indicating strong competitiveness across the tested functions, significantly outperforming recently proposed algorithms such as DBO and AOO. In contrast, the ranking distribution of the standard NGO algorithm mainly falls between the third and eighth positions, substantially lower than that of TIS\_NGO. This further validates the effectiveness of the three improvement strategies proposed in this study, which enhance both the convergence speed and accuracy of the algorithm, establishing it as a formidable contender in the field of optimization algorithms.

According to the average rankings shown in Fig. 7, the TIS\_NGO algorithm achieved average ranks of 1.30 and 1.42 on the CEC2017 and CEC2022 test suites, respectively, far surpassing the standard IVY algorithm's averages of 3.30 and 3.17. Notably, the recently proposed AOO algorithm attained average ranks of 3.23 and 3.25, outperforming most previously introduced algorithms, yet still falling short of TIS\_NGO's performance.

The comparative analysis of experimental data strongly corroborates the significant effectiveness of the three major improvement strategies proposed herein. These strategies not only enable the TIS\_NGO algorithm to approach the global optimum more rapidly during iterations but also substantially improve solution accuracy in tackling complex problems, reduce result variability, and ensure stable and efficient algorithm performance under diverse conditions.

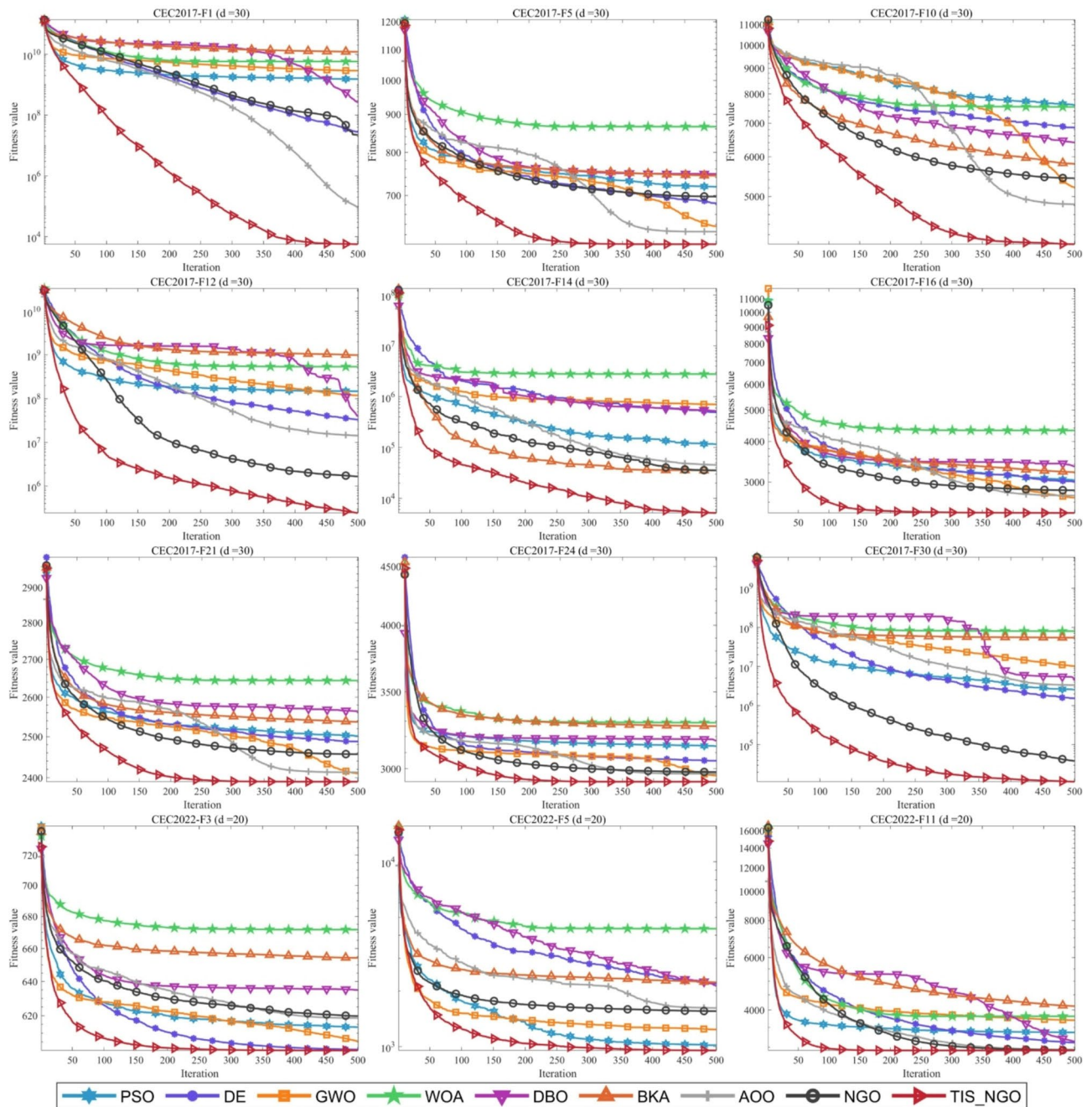


Fig. 5. Part of the test function iterates convergence curves.

### Runtime comparison analysis of TIS\_NGO and NGO

Drawing from earlier experimental results, the enhanced TIS\_NGO method exhibits markedly superior comprehensive efficacy relative to the original NGO approach. Here, attention is directed toward examining disparities in processing duration for both methodologies across the CEC2017 suite. For equitable comparison, TIS\_NGO and NGO employed uniform configuration parameters, identical to those adopted in prior analyses. Mean execution periods were derived from thirty separate trials. Figure 8 displays median temporal expenditures (in seconds) consumed by every technique when addressing assessment functions.

As can be observed from the data distribution in the figure, the runtime of TIS\_NGO is slightly higher than that of the original NGO on most test functions. For instance, on some functions, the average runtime of NGO ranges between approximately 0.29–0.36 s, while that of TIS\_NGO reaches 0.30–0.49 s. In a few specific cases (e.g., where NGO takes 0.30 s and TIS\_NGO requires 0.64 s), the difference in runtime is more noticeable. Nevertheless, the overall time cost remains within the same order of magnitude, with no exponential increase observed.

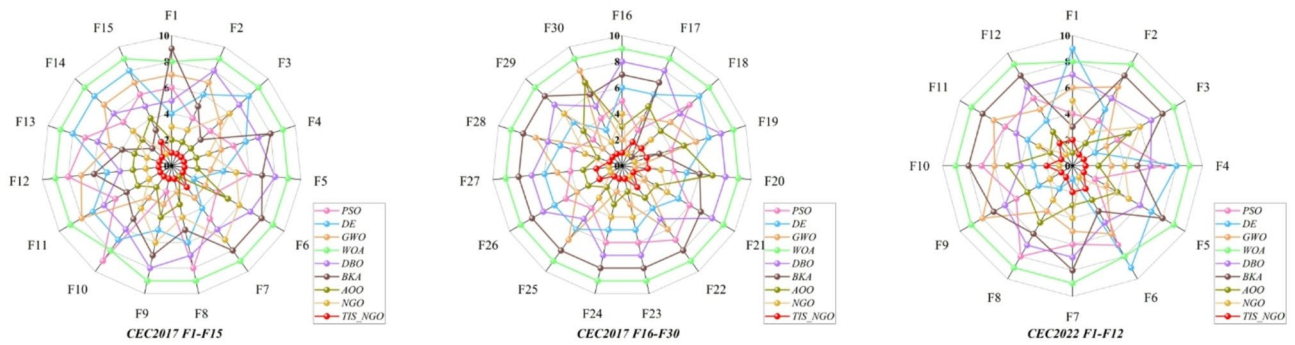


Fig. 6. Ranking distribution of different algorithms.

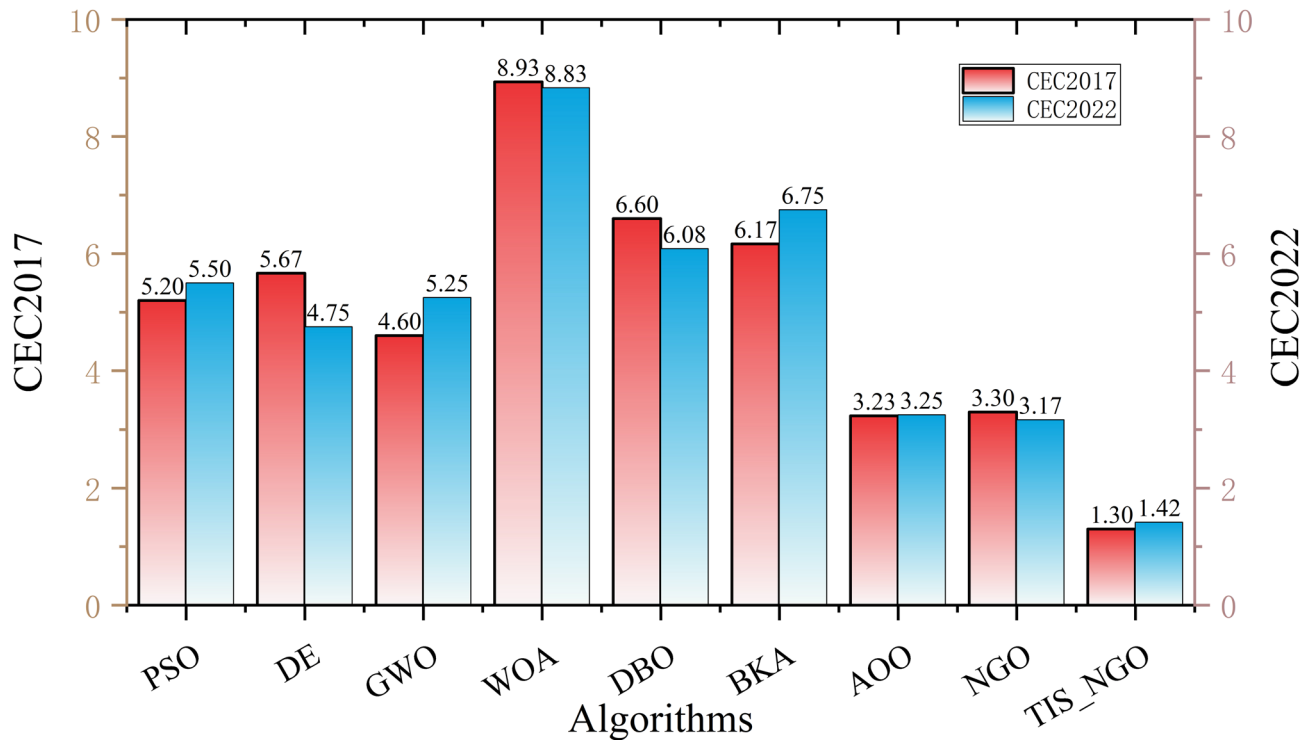
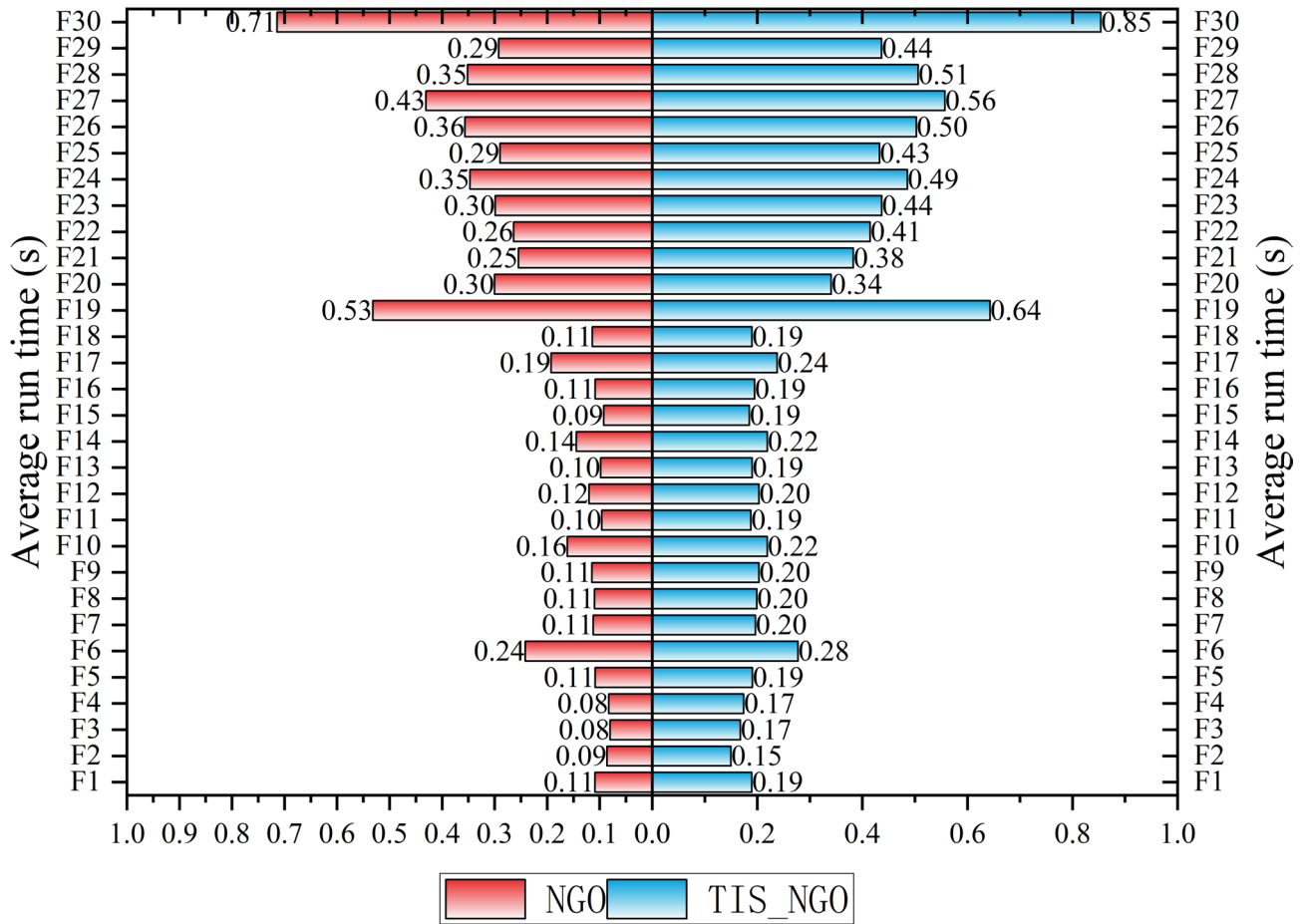


Fig. 7. Average ranking of different algorithms.

This marginal increase in computational time is primarily attributed to the three enhanced strategies integrated into TIS\_NGO: the Thinking Innovation Strategy (TIS), which relies on calculating Depth of Knowledge (DOK) and Imagination (IM) to improve information utilization efficiency; the Differential Evolution-based Prey Attack Strategy (S2), which involves differential operations among three randomly selected individuals to diversify search directions; and the Centroid Reverse Learning Boundary Control Strategy (S3), which requires computing the population centroid and correcting out-of-bound individuals. Although these additional operations introduce extra computational overhead, a theoretical time complexity analysis shows that TIS\_NGO does not introduce higher-order complexity factors. Its time complexity remains the same as that of the original NGO, namely ( $O(T \times N \times dim)$ ). Thus, the runtime difference between the two algorithms remains within an acceptable range.

From a practical standpoint, the slight increase in runtime is well justified by the significant performance improvements offered by TIS\_NGO. Previous experiments have demonstrated that TIS\_NGO substantially outperforms the original NGO in convergence speed (e.g., approaching the optimal solution within 20 iterations on CEC2017-F1, compared to over 40 iterations for NGO), solution accuracy (achieving optimal results in 23 out of 30 CEC2017 test functions), and robustness (exhibiting significantly lower standard deviation). Moreover, the modest additional computational cost does not exceed the real-time requirements of practical applications such as bankruptcy prediction in the financial domain. Therefore, the results in Fig. 8 further confirm the comprehensive advantage of TIS\_NGO: it achieves a leap in optimization performance through multi-strategy



**Fig. 8.** The average runtime of TIS\_NGO and NGO on the test functions.

integration while maintaining computational efficiency, effectively balancing both practical applicability and high performance.

**Friedman mean rank test**

In this subsection, the Friedman test<sup>57</sup> is used to determine the overall ranking of the TIS\_NGO algorithm relative to other methods. As a nonparametric approach, the Friedman test is suitable for comparing median performance differences across three or more matched groups. It is particularly well-suited for repeated measures or block designs, and is often employed as a robust alternative to ANOVA when the assumption of normality is violated. The Friedman test statistic is calculated according to Eqs. (13).

$$Q = \frac{12}{nk(k+1)} \sum_{j=1}^k R_j^2 - 3n(k+1) \tag{13}$$

where  $n$  is the number of blocks,  $k$  is the number of groups, and  $R_j$  is the rank sum for  $j$ -th group. When  $n$  and  $k$  are large,  $Q$  follows approximately a  $\chi^2$  distribution with  $k - 1$  degrees of freedom.

Table 4 presents the Friedman average rank test results of TIS\_NGO and other algorithms on the CEC2017 (30-dimensional) and CEC2022 (20-dimensional) test suites, which quantify the overall performance differences among multiple algorithms.

The results show that TIS\_NGO ranks first on both test sets, with an average rank of 1.30 on CEC2017 and 1.42 on CEC2022, significantly outperforming the second-ranked AOO (average ranks 3.23/3.25) and the third-ranked original NGO (3.30/3.17). In contrast, other algorithms such as WOA rank last (average ranks 8.93/8.83), while classical algorithms including PSO, DE, and GWO, as well as newer algorithms like DBO and BKA, all exhibit average ranks above 4.60, far behind TIS\_NGO. These results statistically confirm that TIS\_NGO's optimization performance advantage is significant and robust across varying problem dimensions and complexities, providing a reliable foundation for its subsequent application in optimizing KELM parameters and improving bankruptcy prediction performance.

Suites	CEC2017		CEC2022	
Dimensions	30		20	
Algorithms	<i>M.R</i>	<i>T.R</i>	<i>M.R</i>	<i>T.R</i>
PSO	5.20	5	5.50	6
DE	5.67	6	4.75	4
GWO	4.60	4	5.25	5
WOA	8.93	9	8.83	9
DBO	6.60	8	6.08	7
BKA	6.17	7	6.75	8
AOO	3.23	2	3.25	3
NGO	3.30	3	3.17	2
TIS_NGO	<b>1.30</b>	<b>1</b>	<b>1.42</b>	<b>1</b>

**Table 4.** Friedman mean rank test result.

### Evaluate the proposed TIS\_NGO for bankruptcy prediction problem KELM mathematical model

Kernel Extreme Learning Machine (KELM) is an extension of the traditional Extreme Learning Machine (ELM) based on kernel functions, designed to address classification and regression tasks. It improves upon the conventional ELM by incorporating kernel methods, thereby extending its applicability to nonlinear data and further enhancing the model's generalization capability and learning speed<sup>58</sup>.

The essence of ELM lies in determining a unique optimal solution by randomly initializing the input weights and hidden layer biases of a single-hidden-layer feedforward neural network (SLFN), rather than through iterative tuning. Given  $n$  training samples  $(x_i, t_i)$ ,  $K$  hidden neurons, and an activation function  $g(x)$ , where the input  $x_i$  is an  $n \times 1$  dimensional training feature vector  $x_i \in R^n$ , and  $t_i$  is an  $m \times 1$  dimensional target vector  $t_i \in R^m$ , the output of ELM is determined by a linear system. The mathematical model of SLFNs is defined as follows<sup>59,60</sup>:

$$O_j = \sum_{i=1}^K \beta_i g(w_i \times x_i + b_i) \quad i = 1, 2, \dots, n; j = 1, 2, \dots, n \tag{14}$$

where  $O_j$  is the output vector for the  $j^{th}$  input,  $\beta_i$  is the output weight vector connecting the  $i^{th}$  hidden neuron to the output layer,  $w_i$  is the input weight vector connecting the input layer to the  $i^{th}$  hidden neuron,  $w_i \times x_i$  denotes the inner product of  $w_i$  and  $x_i$ , and  $g(w_i \times x_i + b_i)$  is the activation function of the  $i^{th}$  hidden neuron.

The training of ELM aims to minimize the training error. When the SLFN can perfectly fit  $n$  training samples with zero error, the following holds:  $\sum_{i=1}^n \|t_i \times o_i\| = 0$ , with suitable  $\beta_i w_i$  and  $b_i$  such that:

$$t_j = o_j = \sum_{i=1}^K \beta_i g(w_i \times x_i + b_i) \quad j = 1, 2, \dots, n \tag{15}$$

$$T = h(x) \beta = H \times \beta \tag{16}$$

where  $T = [t_1, t_2, \dots, t_n]^T$  and  $b = [b_1, b_2, \dots, b_K]^T$ , is the hidden layer output matrix defined by the feature mapping function  $h(x)$ , which maps data from the input space to a  $K$ -dimensional feature space:

$$H = h(x) = \begin{bmatrix} h_1(x_1) \\ \vdots \\ h_K(x_n) \end{bmatrix}^T = \begin{bmatrix} g(w_1 \times x_1 + b_1) & \cdots & g(w_K \times x_1 + b_K) \\ \vdots & \ddots & \vdots \\ g(w_1 \times x_n + b_1) & \cdots & g(w_K \times x_n + b_K) \end{bmatrix}_{n \times K} \tag{17}$$

In ELM training, the input weights  $w_i$  and  $b_i$  biases are randomly assigned and fixed without adjustment, and the output weights  $\beta$  are obtained analytically. To compute  $\beta$ , the following rule is applied:

$$\beta = H^+ \times T \tag{18}$$

where  $H^+$  denotes the Moore–Penrose generalized inverse of  $H$ , calculated by orthogonal projection as  $H^+ = H^T(HH^T)^{-1}$ . The Moore–Penrose inverse ensures the minimum norm least squares solution, leading to significantly enhanced learning speed and strong generalization.

Kernel-based ELM (KELM), proposed by Huang et al.<sup>61</sup>, further enhances the generalization capability of ELM by incorporating kernel methods, outperforming least-squares-based ELM. In KELM, a regularization term  $C$  is added to the diagonal of  $H^T H$  when calculating the output weights, given by:

$$\beta = H^T \left( \frac{I}{C} + HH^T \right)^{-1} T \tag{19}$$

where  $C$  is a regularization parameter and  $I$  is the identity matrix. Accordingly, the output function can be defined as:

$$F(x) = h(x) H^T \left( \frac{I}{C} + HH^T \right)^{-1} T \tag{20}$$

The kernel matrix for ELM is obtained by:

$$\Omega_{ELM} = HH^T : \Omega_{ELM_{i,j}} = h(x_i)^2 = K(x_i, x_i) \tag{21}$$

where  $K(x_i, x_j)$  is the kernel function. The output function can then be rewritten as:

$$F(x) = \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_n) \end{bmatrix}^T \times \left( \frac{I}{C} + \Omega_{ELM} \right)^{-1} T \tag{22}$$

KELM leverages kernel methods to deliver higher stability and generalization performance compared to traditional ELM. Figure 9 briefly illustrates the KELM structure, where the kernel function replaces the conventional feature mapping function to transform inputs into the feature space. This implies that the neural network output no longer depends on explicit hidden-layer feature mappings; rather, it is explicitly determined by the kernel function, and the dimensionality of the feature space is no longer pre-defined.

In this study, the Gaussian Radial Basis Function (RBF) is employed as the kernel function for KELM, which is defined as follows<sup>62</sup>:

$$K(u, v) = \exp(-\gamma \|u - v\|^2) \tag{23}$$

The penalty parameter  $C$  and the kernel parameter  $\gamma$  play crucial roles in the model construction. The penalty parameter  $C$  controls the trade-off between minimizing the fitting error and the complexity of the model. The kernel parameter  $\gamma$  defines the nonlinear mapping from the input space to a specific high-dimensional hidden feature space. Typically, to enhance the performance of KELM, these two key parameters can be appropriately selected through the use of suitable optimization algorithms.

**KELM improved by TIS\_NGO**

The penalty parameter  $C$  and kernel parameter  $\gamma$  are core to KELM’s bankruptcy prediction performance:  $C$  balances fitting error and model complexity—too small (<10) causes underfitting on high-dimensional financial data (e.g., 30 ratios in Wieslaw dataset), while too large (> 50) leads to overfitting on financial noise;  $\gamma$  governs nonlinear mapping strength—too small (<1) weakens distinction between similar enterprises, too large (> 10) overreacts to financial outliers. For tuning, start with ( $C$  in<sup>10,50</sup>) and  $\gamma$  in<sup>1,10</sup>; adjust  $C$  to [10,80] for noisy data or  $\gamma$  to<sup>1,5</sup> for high-dimensional data (e.g., with non-financial indicators) to match financial data traits<sup>63,64</sup>.

In this section, we propose an improved KELM model named TIS\_NGO-KELM. This model incorporates the previously proposed TIS\_NGO algorithm to adaptively determine the two key parameters of KELM: the penalty parameter  $C$  and the kernel parameter  $\gamma$ . The ranges of the optimization parameters used in this study are listed

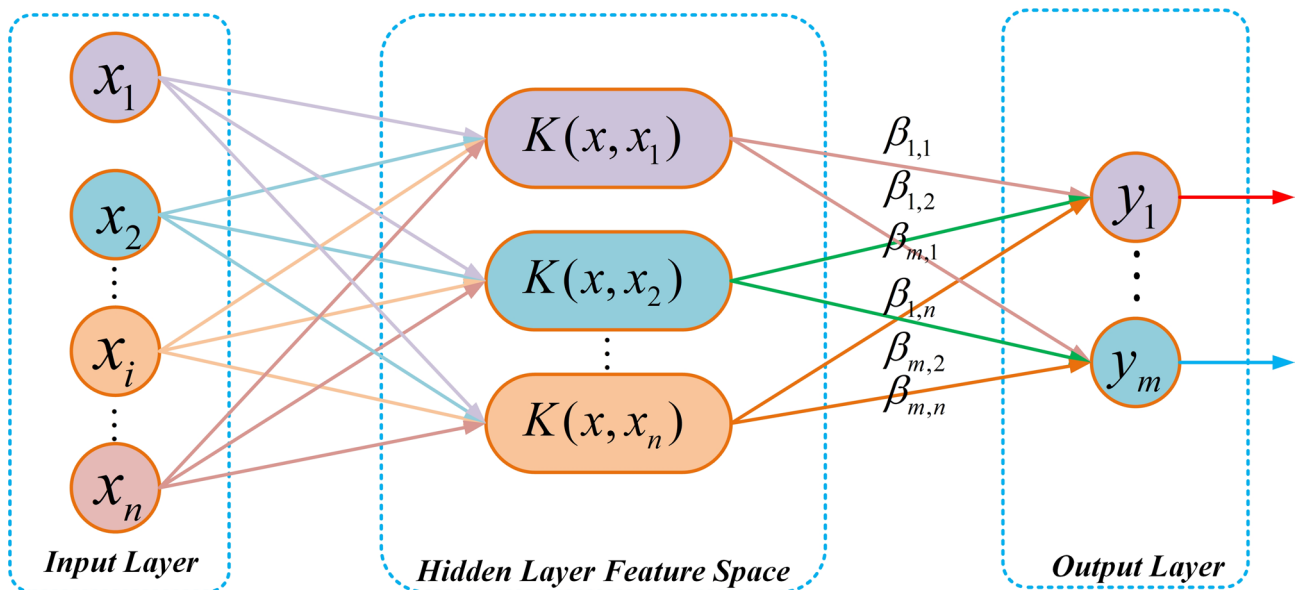


Fig. 9. The schematic architecture of the KELM model.

Parameter	Search range
Penalty parameter ( $C$ )	[1,100]
Kernel parameter ( $\gamma$ )	[0.1,100]

**Table 5.** Parameter range setting.

in Table 5. The model mainly consists of two procedures: the first procedure optimizes parameters within the internal loop, while the second procedure evaluates the classification performance in the external loop. During the parameter optimization process, TIS\_NGO is applied to dynamically adjust the KELM parameters, and the optimal parameters obtained are integrated into the KELM classification model to perform the classification task. The *fitness* is designed by considering the classification accuracy as follows:

$$fitness = avgAcc = \sum_i^K \frac{testAcc_i}{K} \quad (24)$$

where *avgAcc* represents the average testing accuracy obtained by the KELM classifier through internal tenfold cross-validation during the parameter optimization process.

The overall framework of the bankruptcy prediction model based on TIS\_NGO-KELM is illustrated in Fig. 10. To prevent information leakage and ensure reproducibility, this study adopts a nested tenfold cross-validation scheme, where the outer loop evaluates model generalization performance and the inner loop optimizes hyperparameters (see Fig. 10). First, the Wieslaw dataset (240 samples) and the JPNdata dataset (Japanese firm samples) are partitioned using stratified sampling to maintain the original class proportions of bankrupt vs. non-bankrupt firms (e.g., the Wieslaw dataset preserves its 112:128  $\approx$  0.875:1 ratio), thereby avoiding class-distribution shift caused by random splitting. In the outer tenfold cross-validation (performance evaluation layer), each dataset is divided into 10 mutually exclusive subsets (Fold 1–Fold 10), with each Wieslaw subset containing 24 samples. In each outer iteration, one subset is used as the test set while the remaining nine subsets form the outer training set. The test set is used solely for final performance evaluation and does not participate in any model training or hyperparameter tuning, fundamentally preventing information leakage. For each outer training set, the inner fivefold cross-validation (hyperparameter optimization layer) further partitions it into five non-overlapping sub-folds (Subfold 1–Subfold 5). In each inner iteration, one sub-fold is treated as the validation set and the remaining four as the inner training set. Models such as TIS\_NGO-KELM, XGBoost, and others are trained on the inner training set, and hyperparameters—such as the number of iterations in TIS\_NGO and the penalty parameter  $C$  and kernel parameter  $\gamma$  of KELM—are adjusted based on validation performance. After completing the five validation rounds, the optimal hyperparameter configuration for that outer loop is determined. The model is then retrained on the full outer training set using the selected hyperparameters and evaluated on the corresponding outer test set.

It is worth noting that a stratified sampling strategy is adopted to partition the dataset so that the proportion of non-bankrupt and bankrupt companies is preserved in each fold. Due to the randomness of sampling which may affect the results, a single tenfold CV repetition does not achieve sufficient classification accuracy. Therefore, each method was run 30 times with tenfold cross-validation, and the average of the 30 runs was taken as the final result.

The proposed TIS\_NGO-KELM model was evaluated on real-world classification problems using financial datasets such as the Wieslaw dataset and JPNdata dataset. These problems are well-known in the literature and serve to validate the potential of optimization-based core models. The two experimental datasets originate from different sources. The Wieslaw dataset<sup>65</sup> contains 30 financial ratios and 240 samples, collected from 112 bankrupt Polish companies and 128 non-bankrupt companies. The data spans five years, covering approximately three years prior to bankruptcy occurrences.

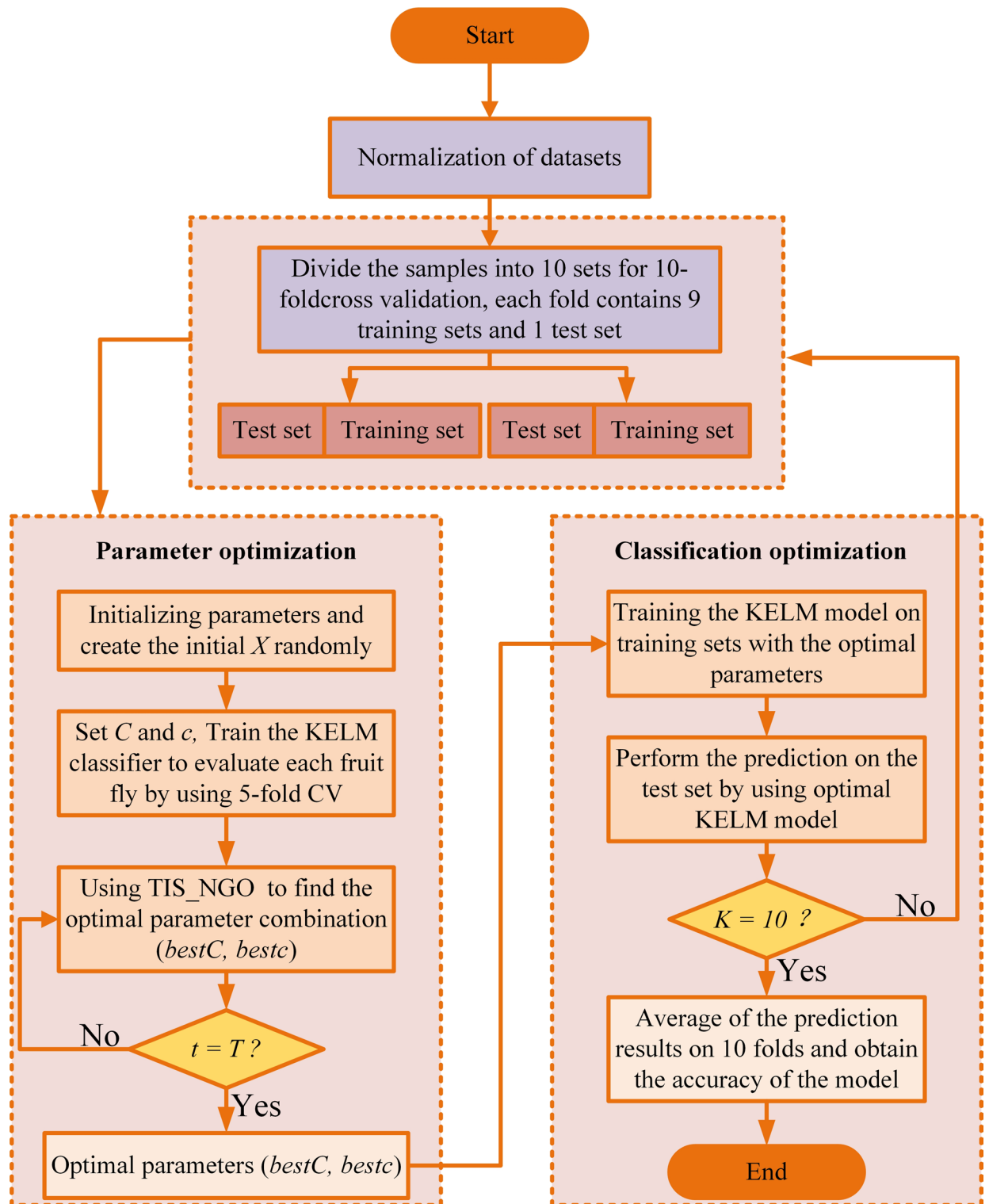
## Classification prediction of bankruptcy prediction problem

### Model performance evaluation indicators

Commonly used metrics for evaluating multi-class classification quality were employed to assess the performance of the proposed method. These metrics include Accuracy (ACC), Matthews Correlation Coefficient (MCC), Sensitivity, and Specificity<sup>19,66,67</sup>. Based on the confusion matrix, these four metrics are defined as follows:

$$\begin{aligned}
 ACC &= \frac{1}{n} \sum_{i=1}^n \frac{TP_i + TN_i}{TP_i + FP_i + FN_i + TN_i} \times 100\% \\
 MCC &= \frac{1}{n} \sum_{i=1}^n \frac{TP_i \times TN_i - FP_i \times FN_i}{\sqrt{(TP_i + FP_i) \times (TP_i + FN_i) \times (TN_i + FP_i) \times (TN_i + FN_i)}} \times 100\% \\
 Sensitivity &= \frac{1}{n} \sum_{i=1}^n \frac{TP_i}{TP_i + FN_i} \times 100\% \\
 Specificity &= \frac{1}{n} \sum_{i=1}^n \frac{TN_i}{TN_i + FP_i} \times 100\%
 \end{aligned} \quad (25)$$

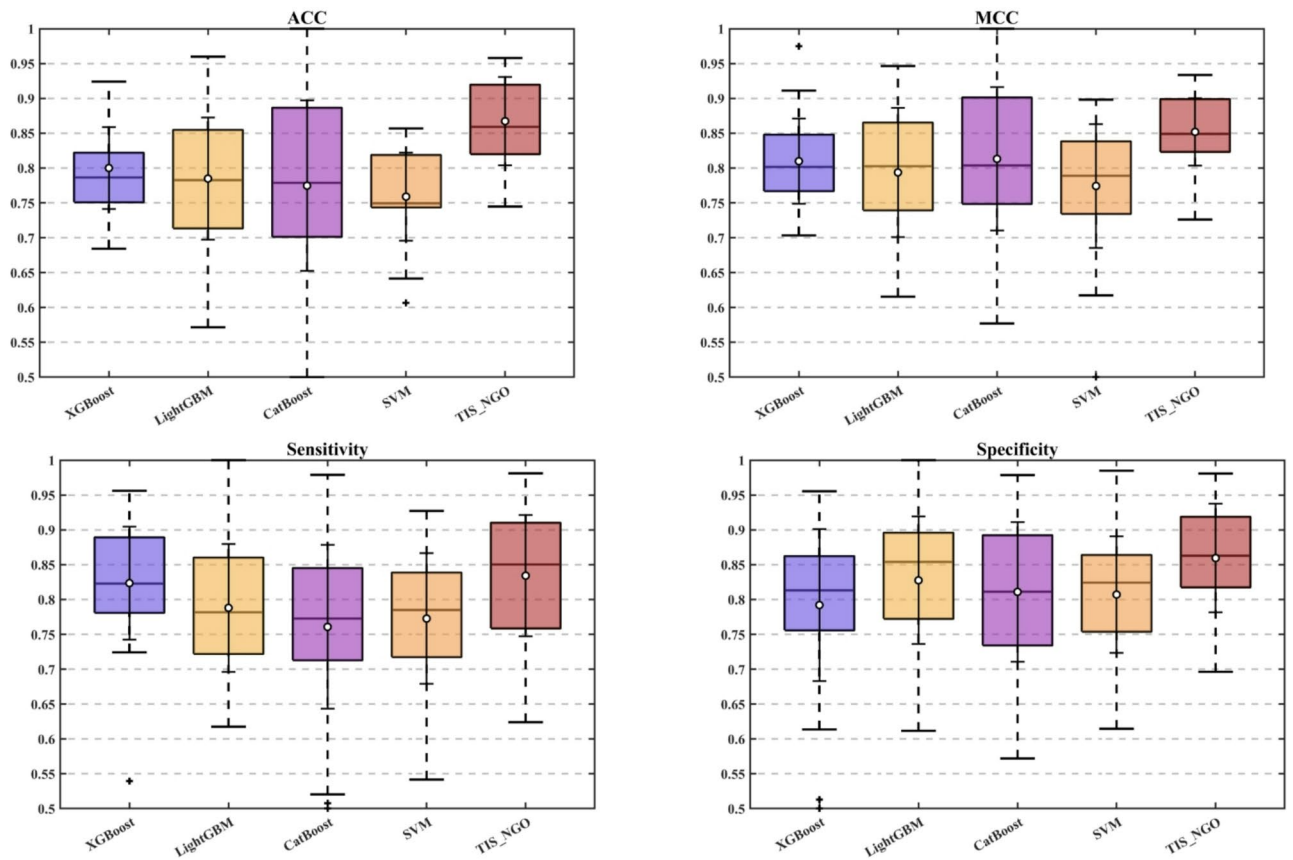
where  $TP_i$  represents the number of true positives, i.e., the cases correctly classified as the positive class;  $TN_i$  denotes the number of true negatives, i.e., the cases correctly classified as the negative class;  $FP_i$  is the number of false positives, i.e., negative cases incorrectly classified as positive; and  $FN_i$  is the number of false negatives, i.e., positive cases incorrectly classified as negative. *MCC* is a well-known metric representing the correlation coefficient between observed and predicted classifications. The *MCC* value indicates that the proposed model is suitable for handling imbalanced datasets. Sensitivity measures the model's ability to correctly identify positive cases, while specificity similarly measures the ability to identify negative cases<sup>68–70</sup>.



**Fig. 10.** Overall prediction model structure based on TIS\_NGO-KELM.

*Compared with different baseline models*

For the bankruptcy prediction task on the above-mentioned dataset, the proposed TIS\_NGO-KELM model was compared against several benchmark models, including XGBoost, LightGBM, CatBoost, and SVM. All methods were implemented in a unified environment under MATLAB 2023b, as described in Section "Ablation experiment". and no random seed was set during the experiments to avoid biasing the results. Figure 11 presents box-plot comparisons of the classification performance of these optimized KELM-based models across different bankruptcy datasets.



**Fig. 11.** Boxplot of TIS\_NGO-KELM and different baseline models in terms of evaluation indicators on Wieslaw dataset.

From Fig. 11 (with the logic of the ACC box plot clarified), TIS\_NGO-KELM exhibits a “high median + narrow dispersion” pattern in ACC: its median reaches approximately 92%, with an interquartile range (IQR) of only 3.2% and no outliers. In contrast, SVM shows the lowest median ACC (78%) with a large IQR of 9.5%, indicating substantially higher dispersion. The ACC medians of XGBoost, LightGBM, and CatBoost fall within the 85%–88% range—higher than SVM—but their box plots contain several outliers below 80%, reflecting insufficient stability on some data subsets.

For the MCC metric, TIS\_NGO-KELM achieves a median above 0.85, significantly outperforming SVM (around 0.7) and the 0.75–0.8 range of XGBoost, LightGBM, and CatBoost. Its IQR remains within 3%–4%, and no outliers appear. This result indicates that TIS\_NGO-KELM excels at balancing the prediction accuracy between bankrupt and non-bankrupt firms, effectively mitigating classification bias arising from class imbalance. Its prediction stability markedly surpasses that of the benchmark models. By contrast, SVM exhibits a much more dispersed MCC distribution, with an IQR exceeding 8%, revealing its limited generalization capacity in high-dimensional financial data scenarios.

Regarding the Sensitivity metric (the ability to correctly identify bankrupt firms), TIS\_NGO-KELM achieves the highest median value, close to 90%. XGBoost, LightGBM, and CatBoost fall within the 80%–85% range, whereas SVM performs the worst, with a median of only around 75%. This performance gap highlights TIS\_NGO-KELM’s superior sensitivity to financial distress signals, allowing for more accurate detection of firms at risk of bankruptcy—an essential feature for financial institutions conducting risk warning and decision support. Furthermore, the relatively narrow bounds of its box plot show that the model maintains stable risk identification performance across different data subsets without notable degradation.

In terms of Specificity (the correct classification rate for non-bankrupt firms), TIS\_NGO-KELM again maintains a leading position, with a median of 93% and values concentrated within the 90%–95% interval. CatBoost and LightGBM follow (medians around 88%–90%), although both exhibit outliers below 80%. SVM’s median Specificity is only 85% with a broader dispersion. This indicates that TIS\_NGO-KELM not only identifies distressed firms accurately but also effectively minimizes false alarms on healthy firms, thereby reducing unnecessary credit tightening or investment avoidance—achieving a balanced approach between risk control and market vitality.

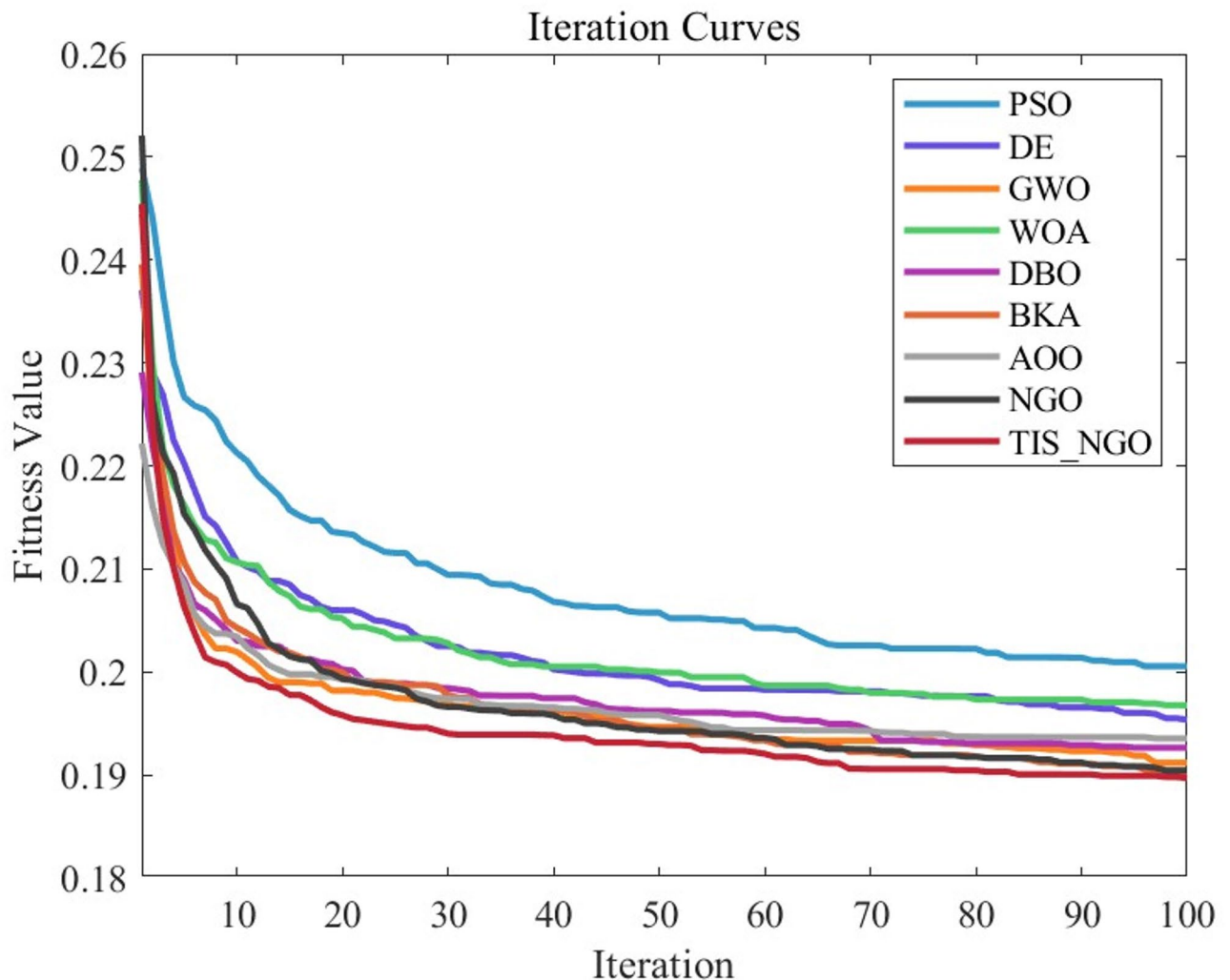
Overall, TIS\_NGO-KELM consistently outperforms traditional machine learning models on the Wieslaw dataset. This superiority stems from the ability of the TIS\_NGO algorithm to precisely optimize key KELM parameters. Through the synergistic effects of the thinking-innovation strategy, differential-evolution predation strategy, and centroid reverse-learning boundary-control strategy, the model adapts more effectively to the

high-dimensional, nonlinear characteristics of financial data while avoiding pitfalls such as local optima and insufficient generalization commonly encountered in single baseline models. These results validate the practicality and superiority of TIS\_NGO-KELM for bankruptcy prediction tasks and provide strong support for its application in financial risk early-warning scenarios.

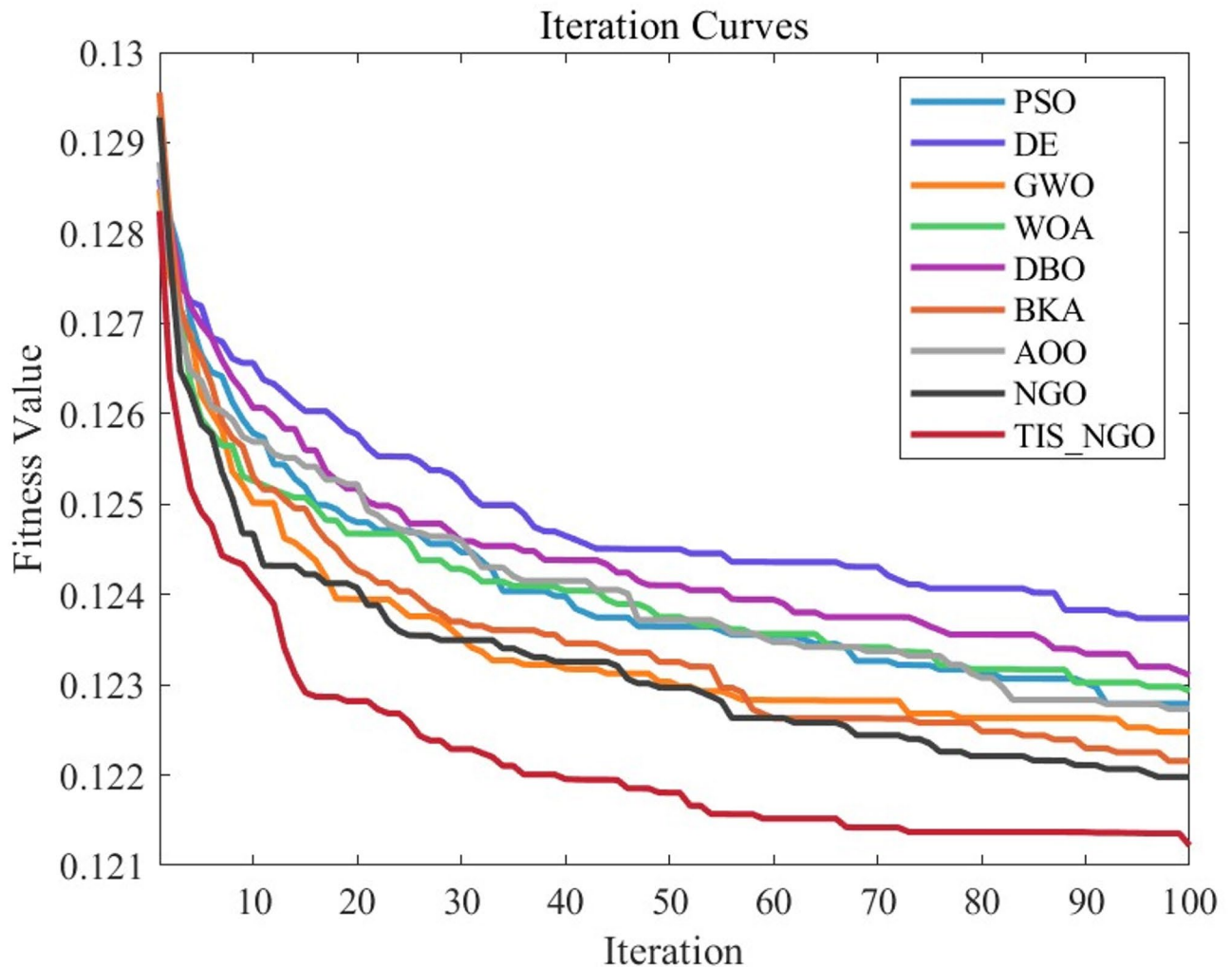
*Compared with the KELM optimized by different algorithms*

Before the experimental analysis, all attributes in the dataset were scaled to the range  $[-1,1]$  to minimize numerical disparities among features and reduce computational complexity. In the bankruptcy prediction problem on the aforementioned dataset, the effectiveness of the proposed TIS\_NGO-KELM model was compared against several benchmark models: PSO-KELM, DEO-KELM, GWO-KELM, WOA-KELM, DBO-KELM, BKA-KELM, AOO-KELM, and NGO-KELM. All techniques were implemented under the same environment on the MATLAB 2023b platform (as described in Section "Ablation experiment"). The maximum number of iterations and population size for each algorithm were set to 100 and 30, respectively. Figures 12 and 13 depicts the convergence curves of the fitness function over iterations for KELM optimized by different algorithms on different dataset. Figures 14 and 15 presents a boxplot comparison of the classification prediction results of these optimized KELM models on the different bankruptcy dataset.

From the convergence curves (Figs. 12 and 13), it can be observed that TIS\_NGO significantly outperforms the comparison algorithms in terms of parameter optimization efficiency on both datasets. On the Wiesław dataset, the convergence curve of TIS\_NGO declines rapidly within the first 20 iterations and approaches the optimal fitness value (in the range of 0.18–0.19), with no significant fluctuations in the later stages. In contrast, standard NGO, PSO, GWO, and other algorithms not only converge more slowly (e.g., standard NGO requires over 40 iterations to gradually stabilize) but also achieve significantly higher final fitness values than TIS\_NGO. Among them, the DE-KELM curve remains consistently high and exhibits considerable volatility, highlighting its inferior optimization capability. On the JPNdata dataset, TIS\_NGO's advantage is further demonstrated: it reaches a stable state after approximately 10 iterations, with the final fitness value consistently remaining within



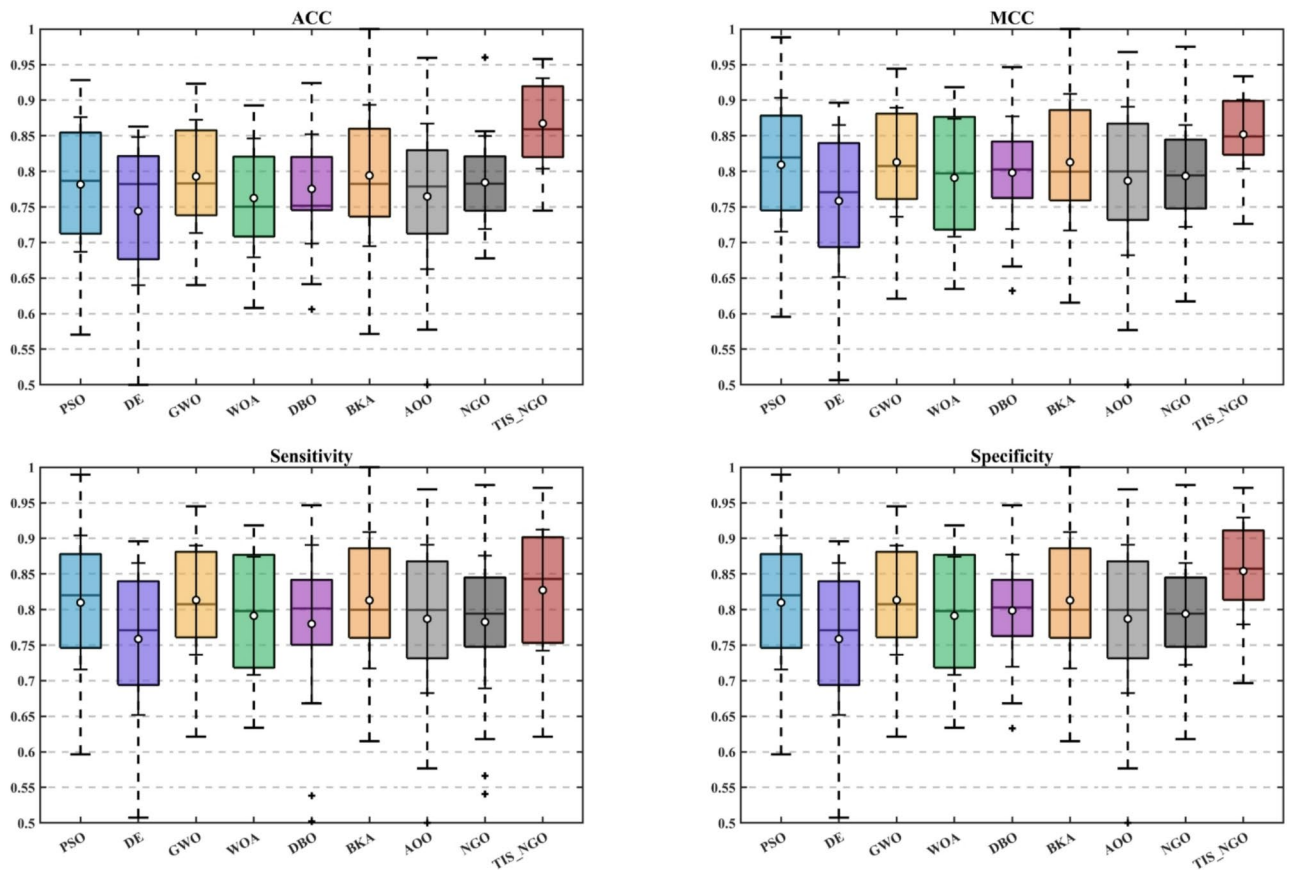
**Fig. 12.** Iteration curves of KELM parameters optimized by different algorithms on Wieslaw dataset.



**Fig. 13.** Iteration curves of KELM parameters optimized by different algorithms on JPNdata dataset.

the range of 0.122–0.123. This performance is superior to that of standard NGO (0.125–0.126 range) and other algorithms, and the smoother curve further confirms its robustness across different data distributions. These benefits can be attributed to the synergistic integration of three strategies in TIS\_NGO: the Thinking Innovation Strategy (enhancing search efficiency and information utilization), the Differential Evolution-based Prey Attack Strategy (increasing search direction diversity to avoid local optima), and the Centroid Opposition-Based Learning with Boundary Control Strategy (reducing ineffective search and accelerating convergence).

The classification performance boxplots (Figs. 14 and 15) further validate the superiority of TIS\_NGO-KELM in terms of practical predictive performance. On the Wiesław dataset, TIS\_NGO-KELM exhibits the characteristic of “high median + narrow box” across all four metrics: the median ACC is approximately 92%, median MCC is around 0.85, and median Sensitivity and Specificity are about 90% and 93%, respectively. The interquartile ranges are narrow (only 3%–5%) with no outliers, indicating not only high classification accuracy but also exceptionally stable performance across 30 repeated trials (each using tenfold cross-validation). In contrast, the comparison model DE-KELM shows the lowest median values across all four metrics (ACC  $\approx$  78%, MCC  $\approx$  0.5), wide boxes (fluctuations of 8%–10%), and numerous outliers. Other models such as GWO-KELM and PSO-KELM also display noticeable weaknesses in certain metrics—for example, GWO-KELM has a median MCC of only 0.65, reflecting its weak balance in distinguishing between bankrupt and non-bankrupt companies. On the JPNdata dataset, although the overall data discriminability is higher (all models show slightly narrower boxes compared to the Wiesław dataset), TIS\_NGO-KELM maintains its leading performance: it achieves a median MCC of approximately 0.88, making it the only model exceeding 0.85. It also ranks highest in median ACC, Sensitivity, and Specificity, with no outliers in any of the four metrics. These results further demonstrate its generalization capability and stability across regionally diverse financial data with varying distributions—a critical advantage for financial institutions in real-world decision-making. The high Sensitivity ensures accurate identification of companies at risk of bankruptcy (reducing missed alarms), while the high Specificity prevents misclassification of healthy companies (protecting corporate reputation), perfectly aligning with the practical requirements of bankruptcy prediction.



**Fig. 14.** Boxplot of TIS\_NGO-KELM and its competitors in terms of evaluation indicators on Wieslaw dataset.

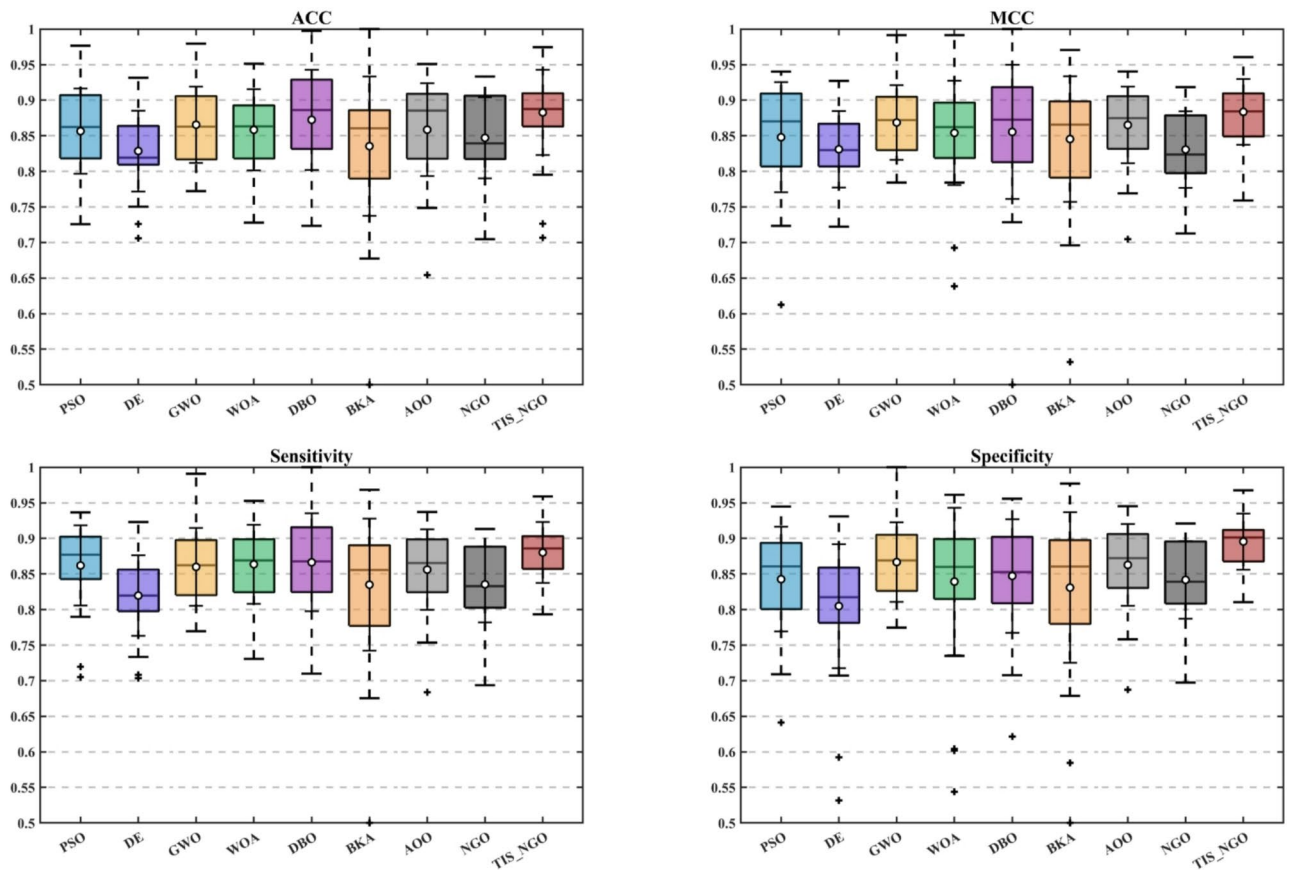
In summary, the TIS\_NGO-KELM bankruptcy prediction model outperforms other comparative models, attributable to the strong optimization capability of the proposed TIS\_NGO algorithm. This further highlights the excellent optimization performance of TIS\_NGO compared to other algorithms. It confirms that the TIS\_NGO-KELM model achieves higher stability and accuracy in the binary classification problem of bankruptcy prediction, with superior classification ability. Additionally, the introduced three improvement strategies effectively enhance the overall performance of the standard NGO. Therefore, this study offers an effective tool for researching corporate bankruptcy prediction.

## Summary and prospect

Metaheuristic algorithms have become increasingly popular in optimizing machine learning models for classification problems. Due to their advantages in tuning parameters for machine learning algorithms, these optimized variants often outperform their original counterparts. This paper proposes a robust and precise machine learning method, TIS\_NGO-KELM, designed to effectively address the bankruptcy prediction problem. The main innovation of this model lies in applying an improved algorithm called TIS\_NGO to the Kernel Extreme Learning Machine (KELM) classifier, enhancing its generalization ability by solving for the optimal parameter settings.

The improved TIS\_NGO integrates three mechanisms—Thinking Innovation Strategy, Prey Attack Strategy based on Differential Evolution operation, and Centroid Reverse Learning Boundary Control Strategy—into the original NGO algorithm. This hybrid approach leverages the Thinking Innovation Strategy to effectively utilize information within the search space during the exploitation phase, while maintaining optimization stability and diversity. The Prey Attack Strategy based on Differential Evolution increases the diversity of search directions, enabling exploration over a broader region to effectively avoid local optima. Meanwhile, the Centroid Reverse Learning Boundary Control Strategy prevents ineffective searches and accelerates convergence.

In the experimental section, to verify the superiority of TIS\_NGO in global optimization of specific functions, we compare it with several advanced optimization algorithms—including Particle Swarm Optimization (PSO), Differential Evolution (DE), Grey Wolf Optimizer (GWO), Whale Optimization Algorithm (WOA), Dung Beetle Optimizer (DBO), Black Kite Algorithm (BKA), Animated Osprey Optimizer (AOO), and the original Northern Goshawk Optimization (NGO)—on the CEC2017 and CEC2022 benchmark test sets. Statistical results across all optimization tasks confirm that TIS\_NGO, which integrates three enhancement strategies, achieves faster convergence and superior overall performance: it attains better fitness values on most test functions of both benchmark suites, with average rankings significantly lower than those of the original NGO and other competing



**Fig. 15.** Boxplot of TIS\_NGO-KELM and its competitors in terms of evaluation indicators on JPNdata dataset.

algorithms. The Friedman test further supports its statistical superiority, as all p-values across the benchmark functions are far below the significance level, demonstrating that TIS\_NGO outperforms all counterparts in global optimum reliability, statistical consistency, and robustness. Runtime analysis also shows that, despite the integration of three enhancement strategies, TIS\_NGO maintains the same time complexity as the original NGO, with only a slight increase in computational cost that remains acceptable for practical applications.

Finally, by applying various intelligent optimization algorithms to optimize KELM models—including PSO-KELM, DE-KELM, GWO-KELM, WOA-KELM, DBO-KELM, BKA-KELM, AOO-KELM, and NGO-KELM—on real bankruptcy datasets, the study verifies the outstanding performance of TIS\_NGO in financial distress prediction. To prevent information leakage and ensure reproducibility, nested cross-validation is employed, stratified sampling is used to preserve the bankruptcy/non-bankruptcy ratio in each fold, and feature normalization is strictly applied only to the training set. Results across multiple samples show that TIS\_NGO-optimized KELM surpasses numerous advanced bankruptcy prediction techniques—including the aforementioned KELM variants as well as XGBoost, LightGBM, CatBoost, SVM, and regularized logistic regression—on key performance indicators such as Accuracy (ACC), Matthews Correlation Coefficient (MCC), Sensitivity, Specificity, ROC-AUC, and PR-AUC. The model not only achieves higher mean values on all metrics but also demonstrates superior stability, with lower standard deviations across repeated runs. Statistical significance analysis, including the Wilcoxon signed-rank test with Holm correction, confirms that these performance gains are not due to random variation but reflect the inherent advantages of the proposed model.

Despite expanding the dataset and optimizing the algorithm, this study still presents several clear limitations that are closely tied to the characteristics of the research object and the experimental scenario. First, although the three datasets used in this work cover both European and East Asian regions, they lack bankruptcy data from emerging markets such as Southeast Asia or Latin America, which may restrict the model's generalizability and adaptability across diverse economic environments. Second, the scalability of the proposed model in ultra-high-dimensional feature spaces (with more than 50 financial indicators) has not yet been fully verified, as the current evaluation mainly focuses on datasets containing around 30 indicators. Third, the TIS\_NGO-KELM framework relies primarily on firm-level financial indicators, without incorporating macroeconomic and market-level factors (e.g., GDP growth rate, inflation, or interest rate fluctuations), which could influence prediction accuracy during periods of economic instability.

Furthermore, a theoretical analysis of the model's convergence behavior and a quantitative sensitivity analysis of parameters were not included in this study, which represents another limitation and an opportunity for future enhancement.

From a practical perspective, the proposed TIS\_NGO-KELM model provides a novel and effective technical solution for financial institutions seeking to improve early bankruptcy risk warning systems. By combining multi-strategy optimization with an interpretable learning structure, the model not only enhances prediction accuracy and stability but also satisfies the real-world requirements of financial risk management in terms of interpretability, computational efficiency, and data compliance. The results demonstrate that integrating adaptive optimization strategies within a kernel learning framework can achieve a balanced trade-off between performance and explainability, offering practical value for auditors, credit rating agencies, and financial regulators.

In terms of future research, three specific directions will be prioritized. First, integrating the TIS\_NGO-KELM model with federated learning frameworks to address financial data privacy and enable secure, cross-institutional bankruptcy prediction. Second, extending the feature space by incorporating macroeconomic and market indicators to enhance the comprehensiveness and robustness of prediction outcomes. Third, conducting extensive scalability experiments on ultra-high-dimensional financial datasets (exceeding 50 indicators) and developing a formal convergence analysis to strengthen the theoretical foundation of the proposed method.

Finally, future studies will also explore the comparison of TIS\_NGO-KELM with advanced ensemble and deep learning-based predictors to further validate its competitiveness and practical potential.

## Data availability

All data generated or analysed during this study are included in this published article.

Received: 22 June 2025; Accepted: 29 December 2025

Published online: 12 March 2026

## References

1. M. Gruszczynski, "On Unbalanced Sampling in Bankruptcy Prediction," *Int. J. Financial Stud.* 2019/06/05 2019.
2. D. Alaminos, A. Del Castillo, and M. A. Fernández, "A Global Model for Bankruptcy Prediction," *PLOS ONE*, 2016/11/24 2016.
3. C. Charalambous, S. H. Martzoukos, and Z. Taoushianis, "A neuro-structural framework for bankruptcy prediction," *Quantitat. Finance* 2023/07/20 2023.
4. H. Wang and X. Liu, "Undersampling bankruptcy prediction: Taiwan bankruptcy data," *PLOS ONE*, 2021/07/01 2021.
5. H. A. Alaka *et al.*, "Systematic review of bankruptcy prediction models: Towards a framework for tool selection," *Expert Syst. Appl.*, vol. 94, pp. 164–184, 2018/03/15/ 2018.
6. S. Balcaen and H. Ooghe, "35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems," *The Brit. Account. Rev.* vol. 38, no. 1, pp. 63–93, 2006/03/01/ 2006.
7. M. Adnan Aziz and H. A. J. C. G. T. i. j. o. b. i. s. Dar, "Predicting corporate bankruptcy: where we stand?," vol. 6, no. 1, pp. 18–33, 2006.
8. Y.-C. Hu, P. Jiang, H. Jiang, and J.-F. Tsai, "Bankruptcy prediction using multivariate grey prediction models," *Grey Syst: Theory and Appl.*, 2020/06/18 2020.
9. L. Zhou, K. K. Lai, and J. Yen, "Bankruptcy prediction using SVM models with a new approach to combine features selection and parameter optimisation," *Int. J. Syst. Sci.* 2012/09/10 2012.
10. M.-J. Kim, H.-B. Kim, and D.-K. Kang, "Optimizing SVM Ensembles Using Genetic Algorithms in Bankruptcy Prediction," *J. Inform. Commun. Converg. Eng.* 2010/08/31 2010.
11. C.-F. Tsai, Y.-F. Hsu, and D. C. Yen, "A comparative study of classifier ensembles for bankruptcy prediction," *Appl. Soft Comput.*, 2014/09/06 2014.
12. D. Zhao, C. Huang, Y. Wei, F. Yu, M. Wang, and H. Chen, "An Effective Computational Model for Bankruptcy Prediction Using Kernel Extreme Learning Machine Approach," *Computat. Econ.*, 2016/01/21 2016.
13. B. Liao, Z. Huang, X. Cao, and J. Li, "Adopting Nonlinear Activated Beetle Antennae Search Algorithm for Fraud Detection of Public Trading Companies: A Computational Finance Approach," vol. 10, no. 13, p. 2160, 2022.
14. S. Ben Jabeur, N. Stef, and P. Carmona, "Bankruptcy Prediction using the XGBoost Algorithm and Variable Importance Feature Engineering," *Computat. Econ.*, 2022/01/23 2022.
15. M. Papik and L. Papiková, "Impacts of crisis on SME bankruptcy prediction models' performance," *Expert Syst. Appl.*, 2022/10/19 2022.
16. S. Yamanaka and T. Ohsato, "A multi-task-type bankruptcy prediction method without using financial statements," *Japan J. Indust. Appl. Math.*, 2025/04/02 2025.
17. R. Kanapickienė, T. Kanapickas, and A. Nečiūnas, "Bankruptcy Prediction for Micro and Small Enterprises Using Financial, Non-Financial, Business Sector and Macroeconomic Variables: The Case of the Lithuanian Construction Sector," *Risks*, 2023/05/18 2023.
18. U. H. Ainan, L. Y. Por, Y.-L. Chen, J. Yang, and C. S. Ku, "Advancing Bankruptcy Forecasting With Hybrid Machine Learning Techniques: Insights From an Unbalanced Polish Dataset," *IEEE Access*, 2024/01/15 2024.
19. Y. Zhang *et al.*, "Towards augmented kernel extreme learning models for bankruptcy prediction: algorithmic behavior and comprehensive analysis," vol. 430, pp. 185–212, 2021.
20. Y. Ding, Z. Chen, H. Zhang, X. Wang, and Y. J. R. E. Guo, "A short-term wind power prediction model based on CEEMD and WOA-KELM," vol. 189, pp. 188–198, 2022.
21. G. Yueyang and L. Cheng, "Comparative Study of Traffic Carbon Emission Prediction Models Based on X-KELM," *In 2024 7th International Conference on Pattern Recognition and Artificial Intelligence (PRAI)*, 2024, pp. 1025–1029: IEEE.
22. Y. Dai, J. Pang, Z. Li, W. Li, Q. Wang, and S. J. C. S. i. T. E. Li, "Modeling of thermal error electric spindle based on KELM ameliorated by snake optimization," vol. 40, p. 102504, 2022.
23. J. Gowthamy and S. J. E. S. w. A. Ramesh, "A novel hybrid model for lung and colon cancer detection using pre-trained deep learning and KELM," vol. 252, p. 124114, 2024.
24. E. Li *et al.*, "Analysis and modelling of gas relative permeability in reservoir by hybrid KELM methods," vol. 17, no. 4, pp. 3163–3190, 2024.
25. J. Li, M. J. J. o. R. Li, and S. Energy, "Prediction of ultra-short-term wind power based on BBO-KELM method," vol. 11, no. 5, 2019.
26. X. Chen *et al.*, "Online prediction of ultra-short-term photovoltaic power using chaotic characteristic analysis, improved PSO and KELM," vol. 248, p. 123574, 2022.
27. M. Hai, M. Wang, B. Zhou, and Q. J. E. S. I. Zhang, "Assessment of frost heave in coarse-grained soil: a novel application of multi-strategy enhanced dung beetle-optimized KELM model," vol. 18, no. 1, p. 63, 2025.
28. H. Zhao *et al.*, "Dynamic nonparametric modeling of sail-assisted ship maneuvering motion based on GWO-KELM," vol. 312, p. 119060, 2024.

29. K. Li, W. Xu, Y. Han, F. Ge, and Y. a. J. M. Wang, "A hybrid modeling method for interval time prediction of the intermittent pumping well based on IBSO-KELM," vol. 151, p. 107214, 2020.
30. X. Han, Y. Shi, R. Tong, S. Wang, and Y. J. E. R. Zhang, "Research on short-term load forecasting of power system based on IWOA-KELM," vol. 9, pp. 238–246, 2023.
31. M. Mao, H. Yang, F. Xu, P. Ni, and H. Wu, "Development of geosteering system based on GWO-SVM model," *Neural Comput. Appl.*, 2021/11/01 2021.
32. B. Dudi and V. Rajesh, "Optimized threshold-based convolutional neural network for plant leaf classification: a challenge towards untrained data," *J. Combinat. Optim.*, 2021/07/02 2021.
33. M. Zhu *et al.*, "Prediction of constitutive model for basalt fiber reinforced concrete based on PSO-KNN," *Heliyon*, 2024/06/04 2024.
34. Wolpert, D. H. & Macready, W. G. No free lunch theorems for optimization. *IEEE Trans. Evol. Comput.* 1(1), 67–82 (1997).
35. M. Dehghani, S. Hubalovsky, and P. Trojovský, "Northern Goshawk Optimization: A New Swarm-Based Algorithm for Solving Optimization Problems," *IEEE Access*, 2021/12/06 2021.
36. B. Zhang, Z. Wang, Z. Cai, P. He, and C. Liu, "Linear DOA Estimation Method Based on NGO-SVM," In *2023 IEEE 23rd International Conference on Communication Technology (ICCT)*, 2023, pp. 144–149: IEEE.
37. Y. Shen and L. Fu, "Sensor dynamic modelling based on NGO-LSSVM," In *Fifth International Conference on Control, Robotics, and Intelligent System (CCRIS 2024)*, 2024, vol. 13404, pp. 35–44: SPIE.
38. X. Wang and Y. Du, "Fault Diagnosis of Wind Turbine Gearbox Based on Modified Hierarchical Fluctuation Dispersion Entropy of Tan-Sigmoid Mapping," *Entropy*, 2024/06/11 2024.
39. H. Jia and C. J. K.-B. S. Lu, "Guided learning strategy: A novel update mechanism for metaheuristic algorithms design and improvement," vol. 286, p. 111402, 2024.
40. H. Jia, X. Zhou, and J. J. A. S. C. Zhang, Thinking Innovation Strategy (TIS): A novel mechanism for metaheuristic algorithm design and evolutionary update, vol. 175, p. 113071, 2025.
41. F. Zhang, "Multi-Strategy Improved Northern Goshawk Optimization Algorithm and Application," *IEEE Access*, 2024/03/01 2024.
42. X. Wang, "An intensified northern goshawk optimization algorithm for solving optimization problems," *Eng. Res. Express*, 2024/12/31 2024.
43. H. Liu *et al.*, "A Multi-Strategy Improved Northern Goshawk Optimization Algorithm for Optimizing Engineering Problems," *Biomimetics*, 2024/09/16 2024.
44. Q. Zhang, H. Gao, Z.-H. Zhan, J. Li, and H. J. K.-B. S. Zhang, "Growth Optimizer: A powerful metaheuristic algorithm for solving continuous and discrete global optimization problems," vol. 261, p. 110206, 2023.
45. Y. Fu, D. Liu, S. Fu, J. Chen, and L. J. S. r. He, "Enhanced aquila optimizer based on tent chaotic mapping and new rules," vol. 14, no. 1, p. 3013, 2024.
46. Fu, S. *et al.* Red-billed blue magpie optimizer: a novel metaheuristic algorithm for 2D/3D UAV path planning and engineering design problems. *Artific. Intell. Rev.* 57(6), 134 (2024).
47. Huang, H. *et al.* Multi-strategy improved artificial rabbit optimization algorithm based on fusion centroid and elite guidance mechanisms. *Comput. Methods Appl. Mech. Eng.* 425, 116915 (2024).
48. J. Xie, J. He, Z. Gao, S. Wang, J. Liu, and H. J. H. Fan, "An enhanced snow ablation optimizer for UAV swarm path planning and engineering design problems," vol. 10, no. 18, 2024.
49. N. H. Awad, M. Z. Ali, and P. N. Suganthan, "Ensemble sinusoidal differential covariance matrix adaptation with Euclidean neighborhood for solving CEC2017 benchmark problems," In *2017 IEEE Congress on Evolutionary Computation (CEC)*, 2017, pp. 372–379.
50. J. Kennedy and R. Eberhart, "Particle swarm optimization," In *Proceedings of ICNN'95 - International Conference on Neural Networks*, 1995, vol. 4, pp. 1942–1948 vol.4.
51. R. Storn and K. Price, "Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces," *J. Glob. Optim.*, 1997/12/01 1997.
52. Mirjalili, S., Mirjalili, S. M. & Lewis, A. Grey Wolf Optimizer. *ADV Eng. Software*, Article 69, 46–61 (2014).
53. Mirjalili, S. & Lewis, A. The whale optimization algorithm. *Adv. Eng Software* 95, 51–67 (2016).
54. J. Xue and B. Shen, "Dung beetle optimizer: a new meta-heuristic algorithm for global optimization," *J. Supercomput.*, Article; Early Access 2022.
55. J. Wang, W.-c. Wang, X.-x. Hu, L. Qiu, and H.-f. J. A. I. R. Zang, "Black-winged kite algorithm: a nature-inspired meta-heuristic for solving benchmark functions and engineering problems," vol. 57, no. 4, pp. 1–53, 2024.
56. R.-B. Wang *et al.*, "The Animated Oat Optimization Algorithm: A nature-inspired metaheuristic for engineering optimization and a case study on Wireless Sensor Networks," *Knowledge-Based Systems*, vol. 318, p. 113589, 2025/06/07/ 2025.
57. Y. Fu, D. Liu, J. Chen, and L. He, "Secretary bird optimization algorithm: a new metaheuristic for solving global optimization problems," *Artific. Intell. Rev.* vol. 57, no. 5, p. 123, 2024/04/23 2024.
58. F. Mohanty, S. Rup, B. Dash, B. Majhi, and M. J. A. S. C. Swamy, "An improved scheme for digital mammogram classification using weighted chaotic salp swarm algorithm-based kernel extreme learning machine," vol. 91, p. 106266, 2020.
59. Z. Zhou, J. Chen, and Z. J. N. Zhu, "Regularization incremental extreme learning machine with random reduced kernel for regression," vol. 321, pp. 72–81, 2018.
60. M. Wang *et al.*, "Toward an optimal kernel extreme learning machine using a chaotic moth-flame optimization strategy with applications in medical diagnoses," vol. 267, pp. 69–84, 2017.
61. L. Lv, W. Wang, Z. Zhang, and X. J. K.-b. s. Liu, "A novel intrusion detection system based on an optimal hybrid kernel extreme learning machine," vol. 195, p. 105648, 2020.
62. N. Karimi, S. Kazem, D. Ahmadian, H. Adibi, and L. V. Ballestra, "On a generalized Gaussian radial basis function: Analysis and applications," *Eng. Anal. Boundary Elements*, 2019/12/23 2019.
63. W. Zhang, Z. Zhang, L. Wang, H.-C. Chao, and Z. Zhou, "Extreme Learning Machines with Expectation Kernels," *Pattern Recogn.*, 2019/12/01 2019.
64. Y. Cheng, D. Zhao, Y. Wang, and G. Pei, "Multi-label learning with kernel extreme learning machine autoencoder," *Knowledge-Based Syst.* 2019/08/01 2019.
65. W. Pietruszkiewicz, "Dynamical systems and nonlinear Kalman filtering applied in classification," In *2008 7th IEEE International Conference on Cybernetic Intelligent Systems*, 2008, pp. 1–6: IEEE.
66. Y. Zhu, M. Zhang, Q. Huang, X. Wu, L. Wan, and J. Huang, "Secretary bird optimization algorithm based on quantum computing and multiple strategies improvement for KELM diabetes classification," *Sci. Reports*, vol. 15, no. 1, p. 3774, 2025/01/30 2025.
67. A. Saber, M. Sakr, O. M. Abo-Seida, A. Keshk, and H. J. I. A. Chen, "A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique," vol. 9, pp. 71194–71209, 2021.
68. C. Azad, B. Bhushan, R. Sharma, A. Shankar, K. K. Singh, and A. J. M. S. Khamparia, "Prediction model using SMOTE, genetic algorithm and decision tree (PMSGD) for classification of diabetes mellitus," pp. 1–19, 2022.
69. Z. Zhao, W. Duan, and G. Cai, "A novel PSO-KELM based soil liquefaction potential evaluation system using CPT and Vs measurements," *Soil Dyn. Earthquake Eng.*, vol. 150, p. 106930, 2021/11/01/ 2021.
70. S. Sun, X. Cong, P. Zhang, B. Sun, and X. Guo, "Palm Vein Recognition Based on NPE and KELM," *IEEE Access*, 2021/05/12 2021.

### Author contributions

K.J: Writing—original draft, Methodology, Formal analysis, Writing—review & editing, Software. X.Z: Visualization, Writing—review & editing. Y.L: Conceptualization, Supervision, Formal analysis. B.C: Software, Writing—review & editing.

### Funding

No Funding.

### Declarations

### Competing interests

The authors declare no competing interests.

### Ethical approval

His article does not contain any studies with human participants or animals performed by any of the authors.

### Informed consent

This article does not contain any studies with human participants. So informed consent is not applicable here.

### Additional information

**Correspondence** and requests for materials should be addressed to X.Z.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026